4 AMPLITUDE MODULATIONS AND DEMODULATIONS

Modulation often refers to a process that moves the message signal into a specific frequency band that is dictated by the physical channel (e.g., voiceband telephone modems). Modulation provides a number of advantages mentioned in Chapter 1, including ease of RF transmission and frequency division multiplexing. Modulations can be analog or digital. Though traditional communication systems such as AM/FM radios and NTSC television signals are based on analog modulations, more recent systems such as second- and third-generation cellular phone systems, HDTV, and DSL are all digital.

In this chapter and the next, we will focus on the classic analog modulations: amplitude modulation and angle modulation. Before we begin our discussion of different analog modulations, it is important to distinguish between communication systems that do not use modulation (**baseband communications**) and systems that use modulation (**carrier communications**).

4.1 BASEBAND VERSUS CARRIER COMMUNICATIONS

The term **baseband** is used to designate the frequency band of the original message signal from the source or the input transducer (see Fig. 1.2). In telephony, the baseband is the audio band (band of voice signals) of 0 to 3.5 kHz. In NTSC television, the video baseband is the video band occupying 0 to 4.3 MHz. For digital data or pulse code modulation that uses bipolar signaling at a rate of R_b pulses per second, the baseband is approximately 0 to R_b Hz.

Baseband Communications

In baseband communication, message signals are directly transmitted without any modification. Because most baseband signals such as audio and video contain significant low-frequency content, they cannot be effectively transmitted over radio (wireless) links. Instead, dedicated user channels such as twisted pairs of copper wires and coaxial cables are assigned to each user for long-distance communications. Because baseband signals have overlapping bands, they would interfere severely if sharing a common channel. Thus, baseband communications leave much of the channel spectrum unused. Modulating several baseband signals and shifting their spectra to nonoverlapping bands allows many users to share one channel by utilizing most of the available bandwidth through frequency division multiplexing (FDM). Long-haul communication over a radio link also requires modulation to shift the signal spectrum to higher frequencies in order to enable efficient power radiation by means of antennas of reasonable dimensions. Yet another use of modulation is to exchange transmission bandwidth for better performance against interferences.

Carrier Modulations

Communication that uses modulation to shift the frequency spectrum of a signal is known as **carrier communication**. In terms of analog modulation, one of the basic parameters (amplitude, frequency, or phase) of a **sinusoidal carrier** of high frequency f_c Hz (or $\omega_c = 2\pi f_c$ rad/s) is varied linearly with the baseband signal $m(t)$. This results in amplitude modulation (AM), frequency modulation (FM), or phase modulation (PM), respectively. Amplitude modulation is linear while the latter two types of carrier modulation are similar and nonlinear, often known collectively as **angle modulation**.

A comment about pulse-modulated signals [pulse amplitude modulation (PAM), pulse width modulation (PWM), pulse position modulation (PPM), pulse code modulation (PCM), and delta modulation (DM)] is in order here. Despite the common term *modulation*, these signals are baseband digital signals. "Modulation" is used here not in the sense of frequency or band shifting. Rather, in these cases it is in fact describing digital pulse-coding schemes used to represent the original analog signals. In other words, the analog message signal is modulating parameters of a digital pulse train. These signals can still modulate a carrier in order to shift their spectra.

Amplitude Modulations and Angle Modulations

We denote as $m(t)$ the source message signal that is to be transmitted by the sender to its receivers and its Fourier transform as $M(f)$. To move the frequency response of $m(t)$ to a new frequency band centered at f_c Hz, the Fourier transform has already revealed a very strong property, the *frequency-shifting* property, that will allow us to achieve this goal. In other words, all we need to do is to multiply $m(t)$ by a sinusoid of frequency f_c such that

$$s_1(t) = m(t)\cos 2\pi f_c t$$

This immediately achieves the basic aim of modulation by moving the signal frequency content to be centered at $\pm f_c$ via

$$S_1(f) = \frac{1}{2}M(f - f_c) + \frac{1}{2}M(f + f_c)$$

This simple multiplication is in fact allowing changes in the amplitude of the sinusoid $s_1(t)$ to be proportional to the message signal. This method is indeed a very valuable modulation known as amplitude modulation.

More broadly, consider a sinusoidal signal

$$s(t) = A(t)\cos[\omega_c t + \phi(t)]$$

There are three variables in a sinusoid: amplitude, (instantaneous) frequency, and phase. Indeed, the message signal can be used to modulate any one of these three parameters to allow $s(t)$ to carry the information from the transmitter to the receiver:

Amplitude $A(t)$ is proportional to $m(t) \Longleftrightarrow$ amplitude modulation
Frequency is proportional to $m(t)$ $\Longleftrightarrow$ frequency modulation
Phase $\phi(t)$ is proportional to $m(t)$ $\Longleftrightarrow$ phase modulation

These processes are known, respectively, as amplitude modulation, frequency modulation, and phase modulation. In this chapter, we describe various forms of amplitude modulations in practical communication systems. Amplitude modulations are linear and their analysis in the time and frequency domains is simpler. In Chapter 5, we will separately discuss the nonlinear angle modulations.

The Interchangeable Use of f and ω

In Chapter 3, we noted the equivalence of frequency response denoted by frequency f or by angular frequency ω. Each of these notations has advantages and disadvantages. After the examples and problems of Chapter 3, readers should be familiar and comfortable with the use of either notation. Thus, from this point on, we use the two different notations interchangeably. Our choice in each case is based on the notational or graphical simplicity.

4.2 DOUBLE-SIDEBAND AMPLITUDE MODULATION

Amplitude modulation is characterized by an information-bearing **carrier** amplitude $A(t)$ that is a linear function of the baseband (message) signal $m(t)$. At the same time, the carrier frequency ω_c and the phase θ_c remain constant. We can assume $\theta_c = 0$ without loss of generality. If the carrier amplitude A is made directly proportional to the modulating signal $m(t)$, then *modulated signal* is $m(t)\cos\omega_c t$ (Fig. 4.1). As we saw earlier [Eq. (3.37)], this type of modulation simply shifts the spectrum of $m(t)$ to the carrier frequency (Fig. 4.1a). Thus, if

$$m(t) \Longleftrightarrow M(f)$$

then

$$m(t)\cos 2\pi f_c t \Longleftrightarrow \frac{1}{2}[M(f+f_c) + M(f-f_c)] \tag{4.1}$$

Recall that $M(f-f_c)$ is $M(f)$ shifted to the right by f_c, and $M(f+f_c)$ is $M(f)$ shifted to the left by f_c. Thus, the process of modulation shifts the spectrum of the modulating signal to the left and to the right by f_c. Note also that if the bandwidth of $m(t)$ is B Hz, then, as seen from Fig. 4.1c, the modulated signal now has bandwidth of $2B$ Hz. We also observe that the modulated signal spectrum centered at $\pm f_c$ (or $\pm\omega_c$ in rad/s) consists of two parts: a portion that lies outside $\pm f_c$, known as the *upper sideband (USB)*, and a portion that lies inside $\pm f_c$, known as the *lower sideband (LSB)*. We can also see from Fig. 4.1c that, unless the message signal $M(f)$ has an impulse at zero frequency, the modulated signal in this scheme does not contain a discrete component of the carrier frequency f_c. In other words, the modulation process does not introduce a sinusoid at f_c. For this reason it is called **double-sideband, suppressed-carrier (DSB-SC) modulation.***

The relationship of B to f_c is of interest. Figure 4.1c shows that $f_c \geq B$, thus avoiding overlap of the modulated spectra centered at f_c and $-f_c$. If $f_c < B$, then the two copies of message spectra overlap and the information of $m(t)$ is distorted during modulation, which makes it impossible to recover $m(t)$ from the modulated signal $m(t)\cos\omega_c t$.

---

* The term *suppressed carrier* does not necessarily mean absence of the spectrum at the carrier frequency f_c. It means that there is no discrete component of the carrier frequency. This implies that the spectrum of the DSB-SC does not have impulses at $\pm f_c$, which also implies that the modulated signal $m(t)\cos 2\pi f_c t$ does not contain a term of the form $k\cos 2\pi f_c t$ [assuming that $m(t)$ has a zero mean value].

Figure 4.1 Double-sideband, suppressed-carrier (DSB-SC) modulation and demodulation.

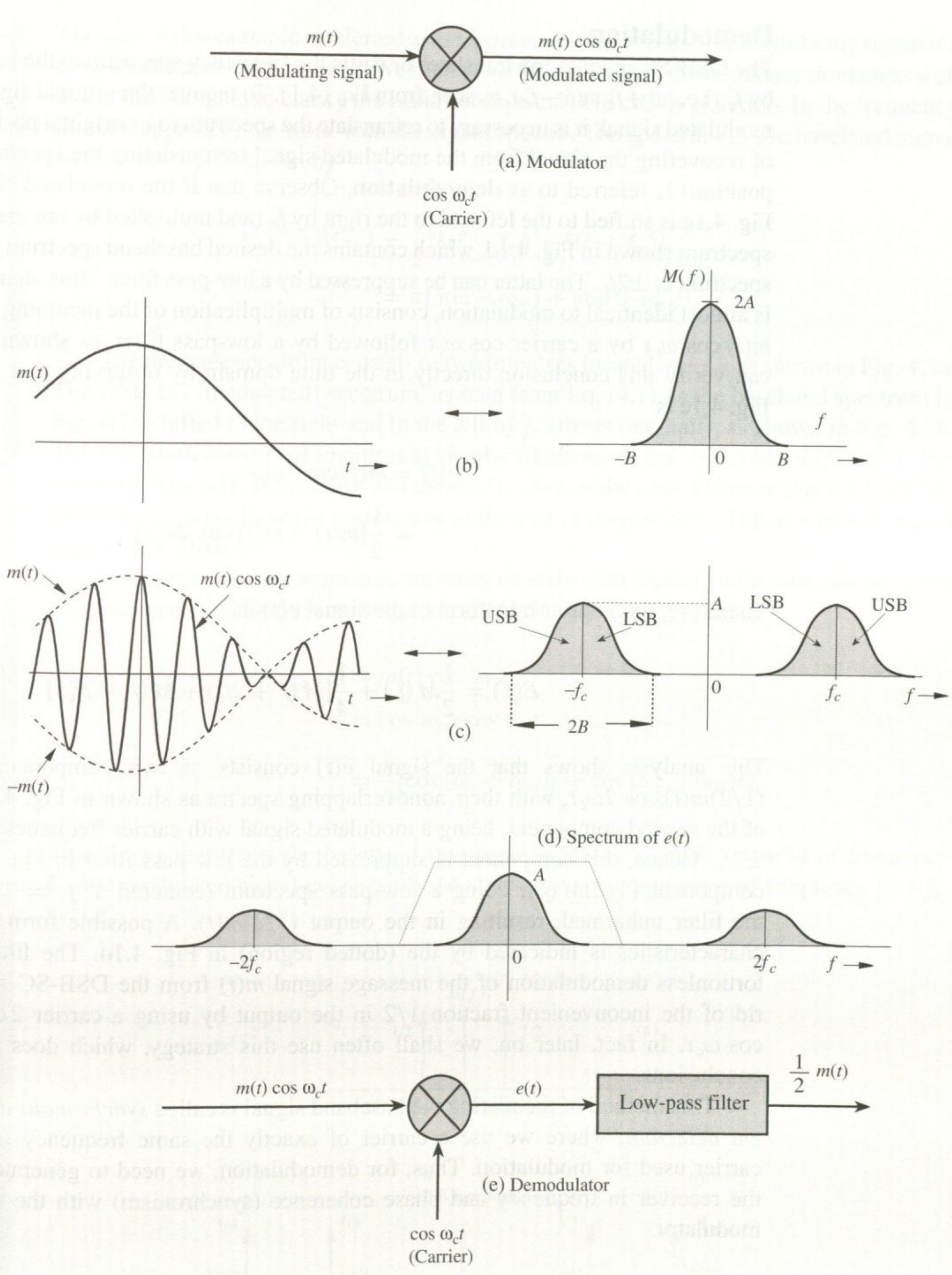

m(t)
(Modulating signal)
m(t) cos ω_c t
(Modulated signal)
(a) Modulator
cos ω_c t
(Carrier)
m(t)
t
(b)
M(f)
2A
f
–B
0
B
m(t)
m(t) cos ω_c t
–m(t)
t
(c)
USB
LSB
A
LSB
USB
–f_c
2B
0
f_c
f
(d) Spectrum of e(t)
A
–2f_c
0
2f_c
f
m(t) cos ω_c t
e(t)
Low-pass filter
1/2 m(t)
(e) Demodulator
cos ω_c t
(Carrier)

Figure 4.1 Double-sideband, suppressed-carrier (DSB-SC) modulation and demodulation.

Note that practical factors may impose additional restrictions on f_c. In broadcast applications, for instance, a transmit antenna can radiate only a narrow band without distortion. This means that to avoid distortion caused by the transmit antenna, we must have $f_c/B \gg 1$. Thus broadcast band AM radio, with $B = 5$ kHz and the band of 550 to 1600 kHz as the carrier frequency, gives a ratio of f_c/B roughly in the range of 100 to 300.

Demodulation

The DSB-SC modulation translates or shifts the frequency spectrum to the left and to the right by f_c (i.e., at $+f_c$ and $-f_c$), as seen from Eq. (4.1). To recover the original signal $m(t)$ from the modulated signal, it is necessary to retranslate the spectrum to its original position. The process of recovering the signal from the modulated signal (retranslating the spectrum to its original position) is referred to as **demodulation**. Observe that if the modulated signal spectrum in Fig. 4.1c is shifted to the left and to the right by f_c (and multiplied by one-half), we obtain the spectrum shown in Fig. 4.1d, which contains the desired baseband spectrum plus an unwanted spectrum at $\pm 2f_c$. The latter can be suppressed by a low-pass filter. Thus, demodulation, which is almost identical to modulation, consists of multiplication of the incoming modulated signal $m(t) \cos \omega_c t$ by a carrier $\cos \omega_c t$ followed by a low-pass filter, as shown in Fig. 4.1e. We can verify this conclusion directly in the time domain by observing that the signal $e(t)$ in Fig. 4.1e is

$$\begin{aligned} e(t) &= m(t) \cos^2 \omega_c t \\ &= \frac{1}{2}[m(t) + m(t) \cos 2\omega_c t] \end{aligned} \tag{4.2a}$$

Therefore, the Fourier transform of the signal $e(t)$ is

$$E(f) = \frac{1}{2}M(f) + \frac{1}{4}[M(f + 2f_c) + M(f - 2f_c)] \tag{4.2b}$$

This analysis shows that the signal $e(t)$ consists of two components $(1/2)m(t)$ and $(1/2)m(t) \cos 2\omega_c t$, with their nonoverlapping spectra as shown in Fig. 4.1d. The spectrum of the second component, being a modulated signal with carrier frequency $2f_c$, is centered at $\pm 2f_c$. Hence, this component is suppressed by the low-pass filter in Fig. 4.1e. The desired component $(1/2)M(f)$, being a low-pass spectrum (centered at $f = 0$), passes through the filter unharmed, resulting in the output $(1/2)m(t)$. A possible form of low-pass filter characteristics is indicated by the (dotted region) in Fig. 4.1d. The filter leads to a distortionless demodulation of the message signal $m(t)$ from the DSB-SC signal. We can get rid of the inconvenient fraction 1/2 in the output by using a carrier $2 \cos \omega_c t$ instead of $\cos \omega_c t$. In fact, later on, we shall often use this strategy, which does not affect general conclusions.

This method of recovering the baseband signal is called *synchronous detection*, or *coherent detection*, where we use a carrier of exactly the same frequency (and phase) as the carrier used for modulation. Thus, for demodulation, we need to generate a local carrier at the receiver in frequency and phase coherence (synchronism) with the carrier used at the modulator.

Example 4.1 For a baseband signal

$$m(t) = \cos \omega_m t = \cos 2\pi f_m t$$

find the DSB-SC signal, and sketch its spectrum. Identify the upper and lower sidebands (USB and LSB). Verify that the DSB-SC modulated signal can be demodulated by the demodulator in Fig. 4.1e.

The case in this example is referred to as *tone modulation* because the modulating signal is a pure sinusoid, or tone, cos $\omega_m t$. We shall work this problem in the frequency domain as well as the time domain to clarify the basic concepts of DSB-SC modulation. In the frequency domain approach, we work with the signal spectra. The spectrum of the baseband signal $m(t) = \cos \omega_m t$ is given by

$$M(f) = \frac{1}{2}[\delta(f - f_m) + \delta(f + f_m)]$$
$$= \pi[\delta(\omega - \omega_m) + \delta(\omega + \omega_m)]$$

The message spectrum consists of two impulses located at $\pm f_m$, as shown in Fig. 4.2a. The DSB-SC (modulated) spectrum, as seen from Eq. (4.1), is the baseband spectrum in Fig. 4.2a shifted to the right and to the left by f_c (times one-half), as shown in Fig. 4.2b. This spectrum consists of impulses at angular frequencies $\pm(f_c - f_m)$ and $\pm(f_c + f_m)$. The spectrum beyond f_c is the USB, and the one below f_c is the LSB. Observe that the DSB-SC spectrum does not have the component of the carrier frequency f_c. This is why it is called *suppressed carrier*.

In the time-domain approach, we work directly with signals in the time domain. For the baseband signal $m(t) = \cos \omega_m t$, the DSB-SC signal $\varphi_{\text{DSB-SC}}(t)$ is

$$\varphi_{\text{DSB-SC}}(t) = m(t) \cos \omega_c t$$
$$= \cos \omega_m t \cos \omega_c t$$
$$= \frac{1}{2}[\cos (\omega_c + \omega_m)t + \cos (\omega_c - \omega_m)t]$$

This shows that when the baseband (message) signal is a single sinusoid of frequency f_m, the modulated signal consists of two sinusoids: the component of frequency $f_c + f_m$

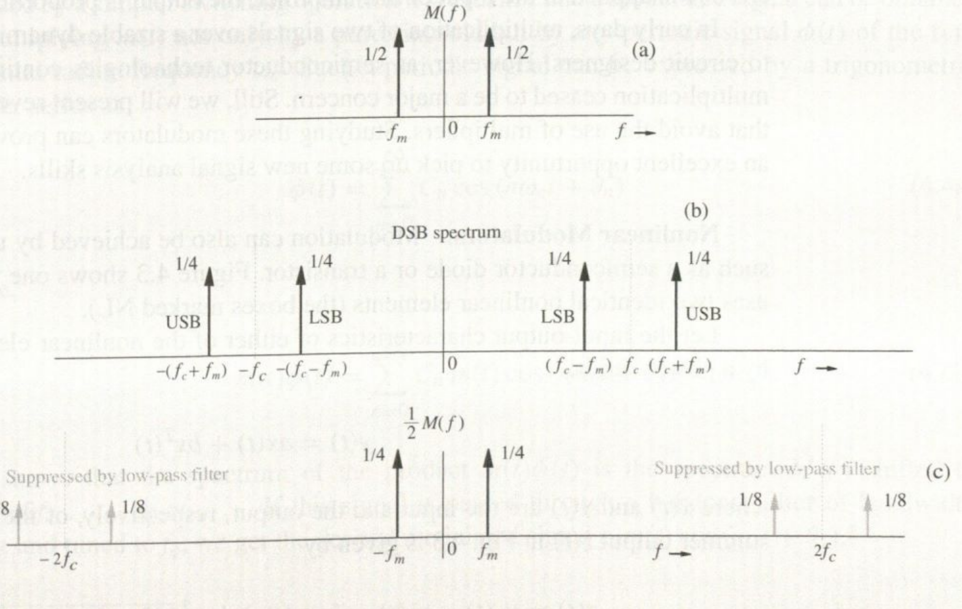

M(f)
1/2
1/2
(a)
$-f_m$
0
f_m
f →
(b)
DSB spectrum
1/4
1/4
1/4
1/4
USB
LSB
LSB
USB
$-(f_c+f_m)$
$-f_c$
$-(f_c-f_m)$
0
(f_c-f_m)
f_c
(f_c+f_m)
f →
$\frac{1}{2}M(f)$
1/4
1/4
Suppressed by low-pass filter
Suppressed by low-pass filter
(c)
1/8
1/8
1/8
1/8
$-2f_c$
$-f_m$
0
f_m
f →
$2f_c$

Figure 4.2
Example of DSB-SC modulation.

(the USB) and the component of frequency $f_c - f_m$ (the LSB). Figure 4.2b shows precisely the spectrum of $\varphi_{\text{DSB-SC}}(t)$. Thus, each component of frequency f_m in the modulating signal turns into two components of frequencies $f_c + f_m$ and $f_c - f_m$ in the modulated signal. Note the fact that there is no component of the carrier frequency f_c on the right-hand side of the preceding equation. As mentioned, this is why it is called double-sideband suppressed-carrier (DSB-SC) modulation.

We now verify that the modulated signal $\varphi_{\text{DSB-SC}}(t) = \cos \omega_m t \cos \omega_c t$, when applied to the input of the demodulator in Fig. 4.1e, yields the output proportional to the desired baseband signal $\cos \omega_m t$. The signal $e(t)$ in Fig. 4.1e is given by

$$\begin{aligned} e(t) &= \cos \omega_m t \cos^2 \omega_c t \\ &= \frac{1}{2} \cos \omega_m t \, (1 + \cos 2\omega_c t) \end{aligned}$$

The spectrum of the term $\cos \omega_m t \cos 2\omega_c t$ is centered at $2f_c$ and will be suppressed by the low-pass filter, yielding $\frac{1}{2} \cos \omega_m t$ as the output. We can also derive this result in the frequency domain. Demodulation causes the spectrum in Fig. 4.2b to shift left and right by f_c (and multiplies by one-half). This results in the spectrum shown in Fig. 4.2c. The low-pass filter suppresses the spectrum centered at $\pm 2f_c$, yielding the spectrum $\frac{1}{2} M(f)$.

Modulators

Modulation can be achieved in several ways. We shall discuss some important categories of modulators.

Multiplier Modulators: Here modulation is achieved directly by using an analog multiplier whose output is proportional to the product of two input signals $m(t)$ and $\cos \omega_c t$. Typically, such a multiplier is obtained from a variable-gain amplifier in which the gain parameter (such as the β of a transistor) is controlled by one of the signals, say, $m(t)$. When the signal $\cos \omega_c t$ is applied at the input of this amplifier, the output is proportional to $m(t) \cos \omega_c t$.

In early days, multiplication of two signals over a sizable dynamic range was a challenge to circuit designers. However, as semiconductor technologies continued to advance, signal multiplication ceased to be a major concern. Still, we will present several classical modulators that avoid the use of multipliers. Studying these modulators can provide unique insights and an excellent opportunity to pick up some new signal analysis skills.

Nonlinear Modulators: Modulation can also be achieved by using nonlinear devices, such as a semiconductor diode or a transistor. Figure 4.3 shows one possible scheme, which uses two identical nonlinear elements (the boxes marked NL).

Figure 4.3 Nonlinear DSB-SC modulator.

Let the input-output characteristics of either of the nonlinear elements be approximated by a power series:

$$y(t) = ax(t) + bx^2(t) \tag{4.3}$$

where $x(t)$ and $y(t)$ are the input and the output, respectively, of the nonlinear element. The summer output $z(t)$ in Fig. 4.3 is given by

$$z(t) = y_1(t) - y_2(t) = [ax_1(t) + b{x_1}^2(t)] - [ax_2(t) + b{x_2}^2(t)]$$

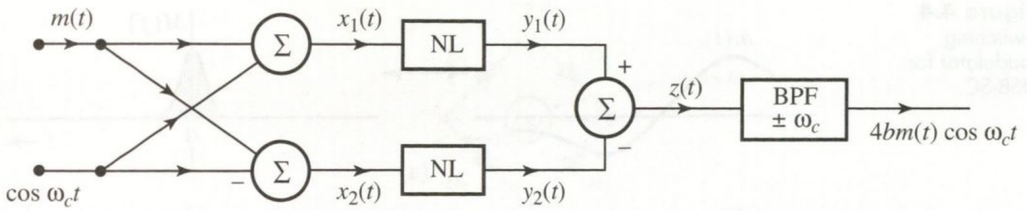


Figure 4.3 Nonlinear DSB-SC modulator.

Substituting the two inputs $x_1(t) = \cos \omega_c t + m(t)$ and $x_2(t) = \cos \omega_c t - m(t)$ in this equation yields

$$z(t) = 2a \cdot m(t) + 4b \cdot m(t) \cos \omega_c t$$

The spectrum of $m(t)$ is centered at the origin, whereas the spectrum of $m(t) \cos \omega_c t$ is centered at $\pm\omega_c$. Consequently, when $z(t)$ is passed through a bandpass filter tuned to ω_c, the signal $am(t)$ is suppressed and the desired modulated signal $4bm(t) \cos \omega_c t$ can pass through the system without distortion.

In this circuit there are two inputs: $m(t)$ and $\cos \omega_c t$. The output of the last summer, $z(t)$, no longer contains one of the inputs, the carrier signal $\cos \omega_c t$. Consequently, the carrier signal does not appear at the input of the final bandpass filter. The circuit acts as a balanced bridge for one of the inputs (the carrier). Circuits that have this characteristic are called *balanced circuits*. The nonlinear modulator in Fig. 4.3 is an example of a class of modulators known as *balanced modulators*. This circuit is balanced with respect to only one input (the carrier); the other input $m(t)$ still appears at the final bandpass filter, which must reject it. For this reason, it is called a *single balanced modulator*. A circuit balanced with respect to both inputs is called a *double balanced modulator*, of which the ring modulator (see later: Fig. 4.6) is an example.

Switching Modulators: The multiplication operation required for modulation can be replaced by a simpler switching operation if we realize that a modulated signal can be obtained by multiplying $m(t)$ not only by a pure sinusoid but by any periodic signal $\phi(t)$ of the fundamental radian frequency ω_c. Such a periodic signal can be expressed by a trigonometric Fourier series as

$$\phi(t) = \sum_{n=0}^{\infty} C_n \cos (n\omega_c t + \theta_n) \tag{4.4a}$$

Hence,

$$m(t)\phi(t) = \sum_{n=0}^{\infty} C_n \, m(t) \cos (n\omega_c t + \theta_n) \tag{4.4b}$$

This shows that the spectrum of the product $m(t)\phi(t)$ is the spectrum $M(f)$ shifted to $\pm f_c,\ \pm 2f_c, \ldots,\ \pm nf_c, \ldots$ If this signal is passed through a bandpass filter of bandwidth $2B$ Hz and tuned to f_c, we get the desired modulated signal $c_1 m(t) \cos (\omega_c t + \theta_1)$.*

---

* The phase θ_1 is not important.

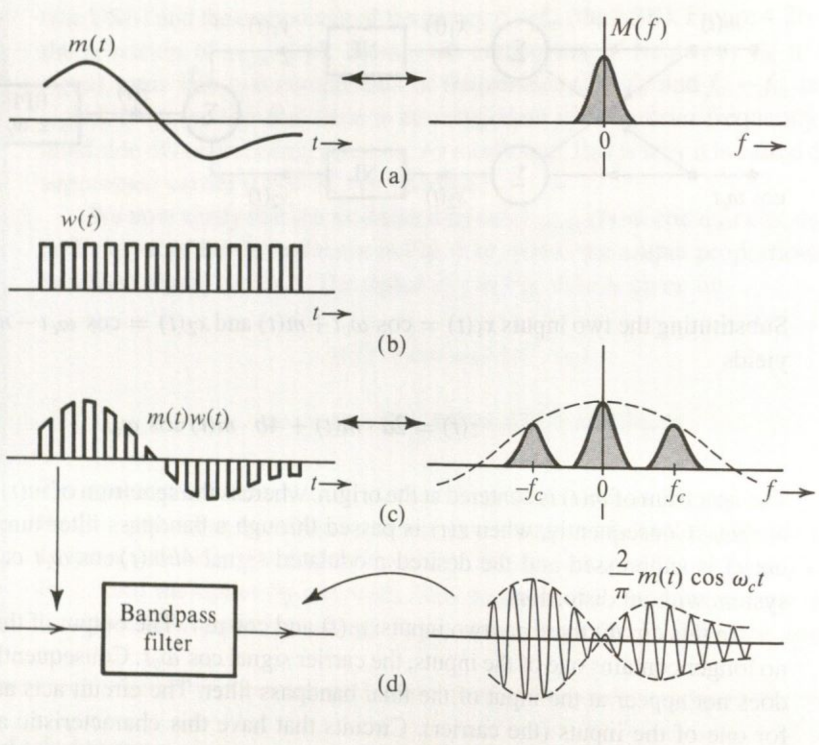


Figure 4.4 Switching modulator for DSB-SC.

The square pulse train $w(t)$ in Fig. 4.4b is a periodic signal whose Fourier series was found earlier in Example 2.8 [Eq. (2.86)] as

$$w(t) = \frac{1}{2} + \frac{2}{\pi}\left(\cos\omega_c t - \frac{1}{3}\cos 3\omega_c t + \frac{1}{5}\cos 5\omega_c t - \cdots\right) \tag{4.5}$$

The signal $m(t)w(t)$ is given by

$$m(t)w(t) = \frac{1}{2}m(t) + \frac{2}{\pi}\left[m(t)\cos\omega_c t - \frac{1}{3}m(t)\cos 3\omega_c t + \frac{1}{5}m(t)\cos 5\omega_c t - \cdots\right] \tag{4.6}$$

The signal $m(t)w(t)$ consists not only of the component $m(t)$ but also of an infinite number of modulated signals with angular frequencies $\omega_c, 3\omega_c, 5\omega_c, \ldots$. Therefore, the spectrum of $m(t)w(t)$ consists of multiple copies of the message spectrum $M(f)$, shifted to $0, \pm f_c, \pm 3f_c, \pm 5f_c, \ldots$ (with decreasing relative weights), as shown in Fig. 4.4c.

For modulation, we are interested in extracting the modulated component $m(t)\cos\omega_c t$ only. To separate this component from the rest of the crowd, we pass the signal $m(t)w(t)$ through a bandpass filter of bandwidth $2B$ Hz (or $4\pi B$ rad/s), centered at the frequency $\pm f_c$. Provided the carrier angular frequency $f_c \geq 2B$ (or $\omega_c \geq 4\pi B$), this will suppress all the spectral components not centered at $\pm f_c$ to yield the desired modulated signal $(2/\pi)m(t)\cos\omega_c t$ (Fig. 4.4d).

We now see the real payoff of this method. Multiplication of a signal by a square pulse train is *in reality* a switching operation. It involves switching the signal $m(t)$ on and off periodically and can be accomplished by simple switching elements controlled by $w(t)$. Figure 4.5a shows one such electronic switch, the **diode bridge modulator,** driven by a sinusoid $A\cos\omega_c t$ to produce the switching action. Diodes D_1, D_2 and D_3, D_4 are matched pairs. When the signal $\cos\omega_c t$ is of a polarity that will make terminal c positive with respect to d, all the diodes

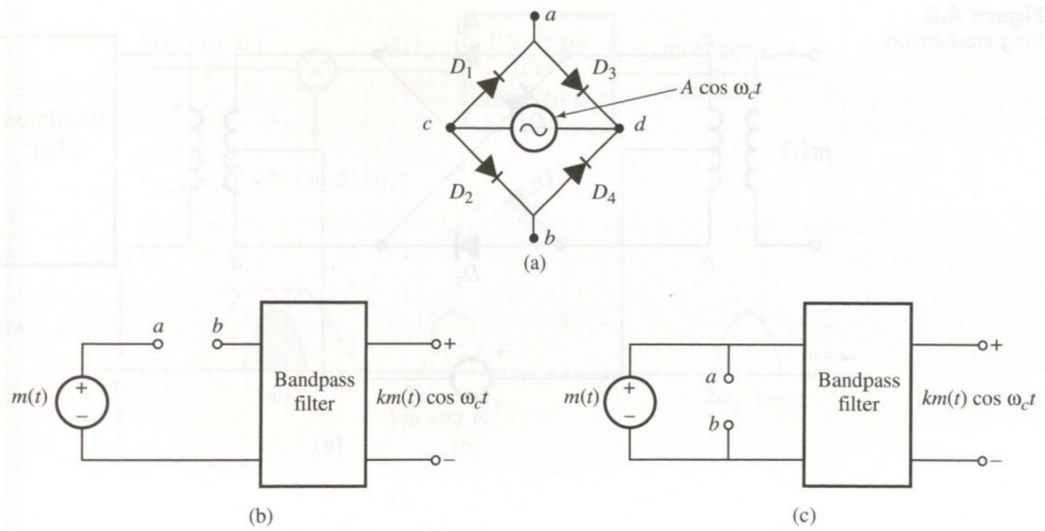


Figure 4.5
(a) Diode-bridge electronic switch.
(b) Series-bridge diode modulator.
(c) Shunt-bridge diode modulator.

conduct. Because diodes D_1 and D_2 are matched, terminals a and b have the same potential and are effectively shorted. During the next half-cycle, terminal d is positive with respect to c, and all four diodes open, thus opening terminals a and b. The diode bridge in Fig. 4.5a, therefore, serves as a desired electronic switch, where terminals a and b open and close periodically with carrier frequency f_c when a sinusoid $A \cos \omega_c t$ is applied across terminals c and d. To obtain the signal $m(t)w(t)$, we may place this electronic switch (terminals a and b) in series (Fig. 4.5b) or across (in parallel) $m(t)$, as shown in Fig. 4.5c. These modulators are known as the **series bridge diode modulator** and the **shunt bridge diode modulator**, respectively. This switching on and off of $m(t)$ repeats for each cycle of the carrier, resulting in the switched signal $m(t)w(t)$, which when bandpass-filtered, yields the desired modulated signal $(2/\pi)m(t) \cos \omega_c t$.

Another switching modulator, known as the **ring modulator**, is shown in Fig. 4.6a. During the positive half-cycles of the carrier, diodes D_1 and D_3 conduct, and D_2 and D_4 are open. Hence, terminal a is connected to c, and terminal b is connected to d. During the negative half-cycles of the carrier, diodes D_1 and D_3 are open, and D_2 and D_4 are conducting, thus connecting terminal a to d and terminal b to c. Hence, the output is proportional to $m(t)$ during the positive half-cycle and to $-m(t)$ during the negative half-cycle. In effect, $m(t)$ is multiplied by a square pulse train $w_0(t)$, as shown in Fig. 4.6b. The Fourier series for $w_0(t)$ as found in Eq. (2.86) is

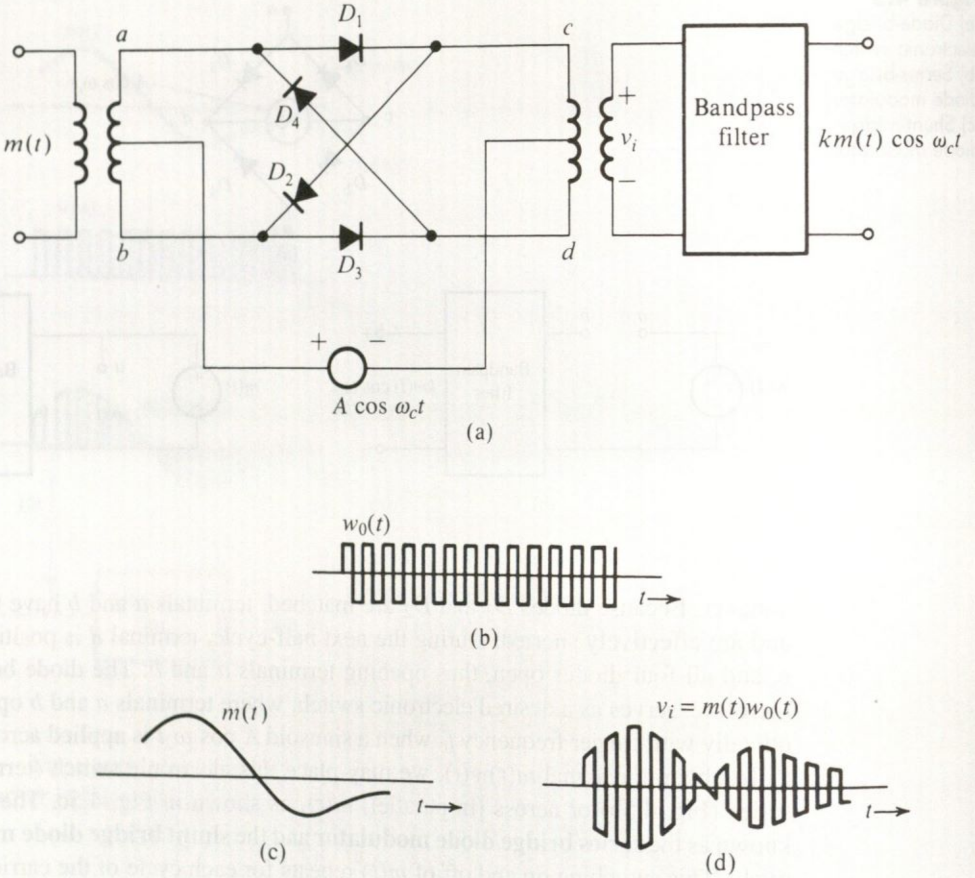
$$w_0(t) = \frac{4}{\pi}\left(\cos \omega_c t - \frac{1}{3}\cos 3\omega_c t + \frac{1}{5}\cos 5\omega_c t - \cdots\right) \tag{4.7a}$$

and

$$v_i(t) = m(t)w_0(t) = \frac{4}{\pi}\left[m(t)\cos \omega_c t - \frac{1}{3}m(t)\cos 3\omega_c t + \frac{1}{5}m(t)\cos 5\omega_c t - \cdots\right] \tag{4.7b}$$

The signal $m(t)w_0(t)$ is shown in Fig. 4.6d. When this waveform is passed through a bandpass filter tuned to ω_c (Fig. 4.6a), the filter output will be the desired signal $(4/\pi)m(t) \cos \omega_c t$.

Figure 4.6
Ring modulation.


a
b
c
d
D_1
D_2
D_3
D_4
$m(t)$
+
v_i
−
Bandpass filter
$km(t) \cos \omega_c t$
+
−
$A \cos \omega_c t$
(a)
$w_0(t)$
$t \rightarrow$
(b)
$m(t)$
$t \rightarrow$
(c)
$v_i = m(t)w_0(t)$
$t \rightarrow$
(d)

In this circuit there are two inputs: $m(t)$ and $\cos \omega_c t$. The input to the final bandpass filter does not contain either of these inputs. Consequently, this circuit is an example of a **double balanced modulator**.

Example 4.2 Frequency Mixer or Converter

We shall analyze a frequency mixer, or frequency converter, used to change the carrier angular frequency of a modulated signal $m(t) \cos \omega_c t$ from ω_c to another frequency ω_I.

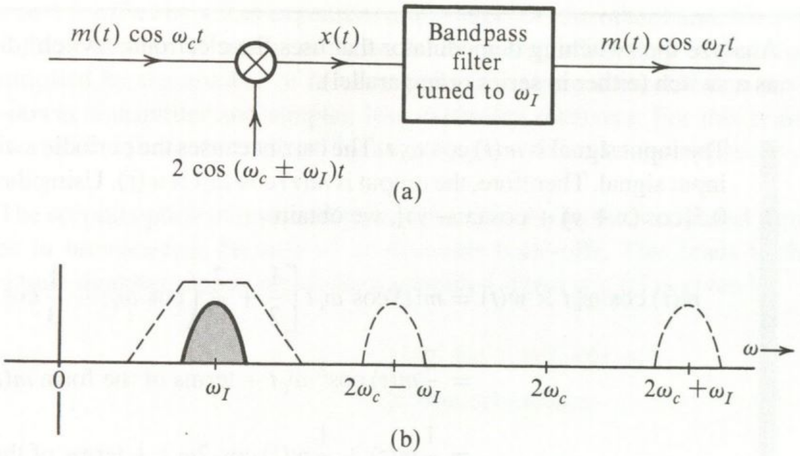
This can be done by multiplying $m(t) \cos \omega_c t$ by $2 \cos \omega_{\text{mix}} t$, where $\omega_{\text{mix}} = \omega_c + \omega_I$ or $\omega_c - \omega_I$, and then bandpass-filtering the product, as shown in Fig. 4.7a. The product $x(t)$ is

$$\begin{aligned} x(t) &= 2m(t) \cos \omega_c t \cos \omega_{\text{mix}} t \\ &= m(t)[\cos (\omega_c - \omega_{\text{mix}})t + \cos (\omega_c + \omega_{\text{mix}})t] \end{aligned}$$

If we select $\omega_{\text{mix}} = \omega_c - \omega_I$, then

$$x(t) = m(t)[\cos \omega_I t + \cos (2\omega_c - \omega_I)t]$$

Figure 4.7
Frequency mixer or converter.



If we select $\omega_{\text{mix}} = \omega_c + \omega_I$, then

$$x(t) = m(t)[\cos\omega_I t + \cos(2\omega_c + \omega_I)t]$$

In either case, as long as $\omega_c - \omega_I \geq 2\pi B$ and $\omega_I \geq 2\pi B$, the various spectra in Fig. 4.7b will not overlap. Consequently, a bandpass filter at the output, tuned to ω_I, will pass the term $m(t)\cos\omega_I t$ and suppress the other term, yielding the output $m(t)\cos\omega_I t$. Thus, the carrier frequency has been translated to ω_I from ω_c.

The operation of frequency mixing/conversion (also known as heterodyning) is basically a shifting of spectra by an additional ω_{mix}. This is equivalent to the operation of modulation with a modulating carrier frequency (the mixer oscillator frequency ω_{mix}) that differs from the incoming carrier frequency by ω_I. Any one of the modulators discussed earlier can be used for frequency mixing. When we select the local carrier frequency $\omega_{\text{mix}} = \omega_c + \omega_I$, the operation is called **super-heterodyning**, and when we select $\omega_{\text{mix}} = \omega_c - \omega_I$, the operation is **sub-heterodyning**.

Demodulation of DSB-SC Signals

As discussed earlier, demodulation of a DSB-SC signal essentially involves a multiplication with the carrier signal and is identical to modulation (see Fig. 4.1). At the receiver, we multiply the incoming signal by a local carrier of frequency and phase in synchronism with the incoming carrier. The product is then passed through a low-pass filter. The **only difference** between the modulator and the demodulator lies in the input signal and the output filter. In the modulator, message $m(t)$ is the input and the multiplier output is passed through a bandpass filter tuned to ω_c, whereas in the demodulator, the DSB-SC signal is the input and the multiplier output is passed through a low-pass filter. Therefore, all the modulators discussed earlier without multipliers can also be used as demodulators, provided the bandpass filters at the output are replaced by low-pass filters of bandwidth B.

For demodulation, the receiver must generate a carrier that is synchronous in phase and in frequency with the incoming carrier. These demodulators are synonymously called **synchronous** or **coherent** (also **homodyne**) demodulators.

Example 4.3 Analyze the switching demodulator that uses the electronic switch (diode bridge) in Fig. 4.5 as a switch (either in series or in parallel).

The input signal is $m(t)\cos\omega_c t$. The carrier causes the periodic switching on and off of the input signal. Therefore, the output is $m(t)\cos\omega_c t \times w(t)$. Using the identity $\cos x\cos y = 0.5[\cos(x+y)+\cos(x-y)]$, we obtain

$$\begin{aligned} m(t)\cos\omega_c t \times w(t) &= m(t)\cos\omega_c t\left[\frac{1}{2}+\frac{2}{\pi}\left(\cos\omega_c t - \frac{1}{3}\cos 3\omega_c t + \cdots\right)\right] \\ &= \frac{2}{\pi}m(t)\cos^2\omega_c t + \text{terms of the form } m(t)\cos n\omega_c t \\ &= \frac{1}{\pi}m(t) + \frac{1}{\pi}m(t)\cos 2\omega_c t + \text{terms of the form } m(t)\cos n\omega_c t \end{aligned}$$

Spectra of the terms of the form $m(t)\cos n\omega_c t$ are centered at $\pm n\omega_c$ rad/s and are filtered out by the low-pass filter yielding the output $(1/\pi)m(t)$. It is left as an exercise for the reader to show that the output of the ring demodulator in Fig. 4.6a (with the low-pass filter at the output) is $(2/\pi)m(t)$ (twice that of the switching demodulator in this example).

4.3 AMPLITUDE MODULATION (AM)

In the last section, we began our discussion of amplitude modulation by introducing DSB-SC amplitude modulation because it is easy to understand and to analyze in both time and frequency domains. However, analytical simplicity is not always accompanied by an equivalent simplicity in practical implementation. The (coherent) demodulation of DSB-SC signal requires the receiver to possess a carrier signal that is synchronized with the incoming carrier. This requirement is not easy to achieve in practice. Because the modulated signal may have traveled hundreds of miles and could even suffer from some unknown frequency shift, the bandpass received signal in fact has the form of

$$r(t) = A_c m(t-t_0)\cos[(\omega_c+\Delta\omega)(t-t_0)] = A_c m(t-t_0)\cos[(\omega_c+\Delta\omega)t-\theta_d)]$$

in which $\Delta\omega$ represents the Doppler effect while

$$\theta_d = (\omega_c+\Delta\omega)t_d$$

comes from the unknown delay t_0. To utilize the coherent demodulator, the receiver must be sophisticated enough to generate a local oscillator $\cos[(\omega_c+\Delta\omega)t-\theta_d)]$ purely from the received signal $r(t)$. Such a receiver would be harder to implement and could be quite costly. This cost should be particularly avoided in broadcasting systems, which have many receivers for every transmitter.

The alternative to a coherent demodulator is for the transmitter to send a carrier $A\cos\omega_c t$ [along with the modulated signal $m(t)\cos\omega_c t$] so that there is no need to generate a carrier at the receiver. In this case the transmitter needs to transmit at a much higher power level, which increases its cost as a trade-off. In point-to-point communications, where there is one transmitter for every receiver, substantial complexity in the receiver system can be justified if the cost is offset by a less expensive transmitter. On the other hand, for a broadcast system with a huge number of receivers for each transmitter, where any cost saving at the receiver is multiplied by the number of receiver units, it is more economical to have one expensive high-power transmitter and simpler, less expensive receivers. For this reason, broadcasting systems tend to favor the trade-off by migrating cost from the (many) receivers to the (fewer) transmitters.

The second option of transmitting a carrier along with the modulated signal is the obvious choice in broadcasting because of its desirable trade-offs. This leads to the so-called AM (amplitude modulation), in which the transmitted signal $\varphi_{\text{AM}}(t)$ is given by

$$\varphi_{\text{AM}}(t) = A\cos\omega_c t + m(t)\cos\omega_c t \tag{4.8a}$$

$$= [A + m(t)]\cos\omega_c t \tag{4.8b}$$

The spectrum of $\varphi_{\text{AM}}(t)$ is basically the same as that of $\varphi_{\text{DSB-SC}}(t) = m(t)\cos\omega_c t$ except for the two additional impulses at $\pm f_c$,

$$\varphi_{\text{AM}}(t) \Longleftrightarrow \frac{1}{2}[M(f+f_c) + M(f-f_c)] + \frac{A}{2}[\delta(f+f_c) + \delta(f-f_c)] \tag{4.8c}$$

Upon comparing $\varphi_{\text{AM}}(t)$ with $\varphi_{\text{DSB-SC}}(t) = m(t)\cos\omega_c t$, it is clear that the AM signal is identical to the DSB-SC signal with $A + m(t)$ as the modulating signal [instead of $m(t)$]. The value of A is always chosen to be positive. Therefore, to sketch $\varphi_{\text{AM}}(t)$, we sketch the envelope $|A+m(t)|$ and its mirror image $-|A+m(t)|$ and fill in between with the sinusoid of the carrier frequency f_c. The size of A affects the time domain envelope of the modulated signal.

Two cases are considered in Fig. 4.8. In the first case, A is large enough to ensure that $A + m(t) \geq 0$ is always nonnegative. In the second case, A is not large enough to satisfy this condition. In the first case, the envelope has the same shape as $m(t)$ (although riding on a direct current of magnitude A). In the second case, the envelope shape differs from the shape of $m(t)$ because the negative part of $A + m(t)$ is rectified. This means we can detect the desired signal $m(t)$ by detecting the envelope in the first case when $A + m(t) > 0$. Such detection is not possible in the second case. We shall see that envelope detection is an extremely simple and inexpensive operation, which does not require generation of a local carrier for the demodulation. But we have shown that the envelope of AM has the information about $m(t)$ only if the AM signal $[A + m(t)]\cos\omega_c t$ satisfies the condition $A + m(t) > 0$ for all t.

Let us now be more precise about the definition of "envelope." Consider a signal $E(t)\cos\omega_c t$. If $E(t)$ varies slowly in comparison with the sinusoidal carrier $\cos\omega_c t$, then the **envelope** of $E(t)\cos\omega_c t$ is $|E(t)|$. This means [see Eq. (4.8b)] that if and only if $A+m(t) \geq 0$ for all t, the envelope of $\varphi_{\text{AM}}(t)$ is

$$|A + m(t)| = A + m(t)$$

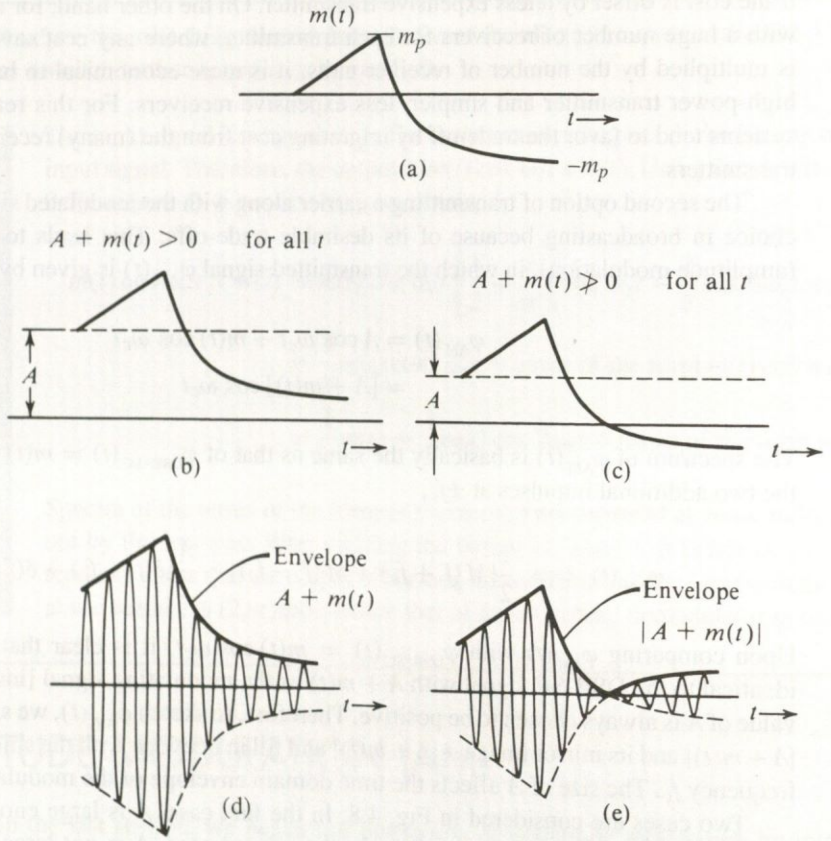
In other words, for envelope detection to properly detect $m(t)$, two conditions must be met:

(a) $f_c \gg$ bandwidth of $m(t)$

(b) $A + m(t) \geq 0$

This conclusion is readily verified from Fig. 4.8d and e. In Fig. 4.8d, where $A + m(t) \geq 0$, $A + m(t)$ is indeed the envelope, and $m(t)$ can be recovered from this envelope. In Fig. 4.8e, where $A + m(t)$ is not always positive, the envelope $|A + m(t)|$ is rectified from $A + m(t)$,

Figure 4.8
AM signal and its envelope.


$m(t)$
m_p
$t \rightarrow$
(a)
$-m_p$
$A + m(t) > 0$ for all t
$A + m(t) \ngtr 0$ for all t
A
$t \rightarrow$
(b)
A
$t \rightarrow$
(c)
Envelope
$A + m(t)$
$t \rightarrow$
(d)
Envelope
$|A + m(t)|$
$t \rightarrow$
(e)

and $m(t)$ cannot be recovered from the envelope. Consequently, demodulation of $\varphi_{\text{AM}}(t)$ in Fig. 4.8d amounts to simple envelope detection. Thus, the **condition for envelope detection** of an AM signal is

$$A + m(t) \geq 0 \qquad \text{for all } t \tag{4.9a}$$

If $m(t) \geq 0$ for all t, then $A = 0$ already satisfies condition (4.9a). In this case there is no need to add any carrier because the envelope of the DSB-SC signal $m(t) \cos \omega_c t$ is $m(t)$, and such a DSB-SC signal can be detected by envelope detection. In the following discussion we assume that $m(t) \ngeq 0$ for all t; that is, $m(t)$ can be negative over some range of t.

Message Signals $m(t)$ with Zero Offset: Let $\pm m_p$ be the maximum and the minimum values of $m(t)$, respectively (Fig. 4.8). This means that $m(t) \geq -m_p$. Hence, the condition of envelope detection (4.9a) is equivalent to

$$A \geq -m_{\min} = m_p \tag{4.9b}$$

Thus, the minimum carrier amplitude required for the viability of envelope detection is m_p. This is quite clear from Fig. 4.8. We define the modulation index μ as

$$\mu = \frac{m_p}{A} \tag{4.10a}$$

For envelope detection to be distortionless, the condition is $A \geq m_p$. Hence, it follows that

$$0 \leq \mu \leq 1 \tag{4.10b}$$

is the required condition for the distortionless demodulation of AM by an envelope detector.

When $A < m_p$, Eq. (4.10a) shows that $\mu > 1$ (overmodulation). In this case, the option of envelope detection is no longer viable. We then need to use synchronous demodulation. Note that synchronous demodulation can be used for any value of μ, since the demodulator will recover signal $A + m(t)$. Only an additional dc blocker is needed to remove the dc voltage A. The envelope detector, which is considerably simpler and less expensive than the synchronous detector, can be used only for $\mu \leq 1$.

Message Signals $m(t)$ with Nonzero Offset: On rare occasions, the message signal $m(t)$ will have a nonzero offset such that its maximum $m_{\max}$ and its minimum $m_{\min}$ are not symmetric, that is,

$$m_{\min} \neq -m_{\max}$$

In such a case, it can be recognized that any offset to the envelope does not change the shape of the envelope detector output. In fact, one should note that constant offset does not carry any fresh information. In this case, envelope detection would still remain distortionless if

$$0 \leq \mu \leq 1 \tag{4.11a}$$

with a modified modulation index definition of

$$\mu = \frac{m_{\max} - m_{\min}}{2A + m_{\max} + m_{\min}} \tag{4.11b}$$

Example 4.4 Sketch $\varphi_{\text{AM}}(t)$ for modulation indices of $\mu = 0.5$ and $\mu = 1$, when $m(t) = b \cos \omega_m t$. This case is referred to as **tone modulation** because the modulating signal is a pure sinusoid (or tone).

In this case, $m_{\max} = b$ and $m_{\min} = -b$. Hence the modulation index according to Eq. (4.10a) is

$$\mu = \frac{b}{A}$$

Hence, $b = \mu A$ and

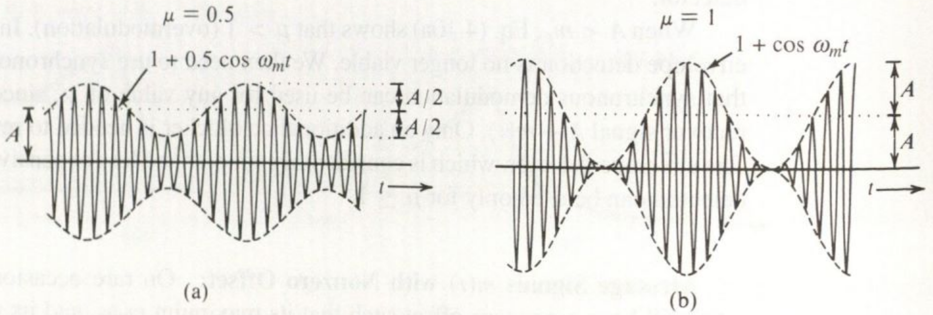
$$m(t) = b \cos \omega_m t = \mu A \cos \omega_m t$$

Therefore,

$$\varphi_{\text{AM}}(t) = [A + m(t)]\cos\omega_c t = A[1 + \mu\cos\omega_m t]\cos\omega_c t$$

Figure 4.9 shows the modulated signals corresponding to $\mu = 0.5$ and $\mu = 1$, respectively.

Figure 4.9 Tone-modulated AM: (a) $\mu = 0.5$; (b) $\mu = 1$.



Sideband and Carrier Power

The advantage of envelope detection in AM has its price. In AM, the carrier term does not carry any information, and hence, the carrier power is wasteful from this point of view,

$$\varphi_{\text{AM}}(t) = \underbrace{A\cos\omega_c t}_{\text{carrier}} + \underbrace{m(t)\cos\omega_c t}_{\text{sidebands}}$$

The carrier power P_c is the mean square value of $A\cos\omega_c t$, which is $A^2/2$. The sideband power P_s is the power of $m(t)\cos\omega_c t$, which is $0.5\,\widetilde{m^2(t)}$ [see Eq. (3.94)]. Hence,

$$P_c = \frac{A^2}{2} \quad\text{and}\quad P_s = \frac{1}{2}\,\widetilde{m^2(t)}$$

The useful message information resides in the sideband power, whereas the carrier power is used for convenience in modulation and demodulation. The total power is the sum of the carrier (wasted) power and the sideband (useful) power. Hence, η, the power efficiency, is

$$\eta = \frac{\text{useful power}}{\text{total power}} = \frac{P_s}{P_c + P_s} = \frac{\widetilde{m^2(t)}}{A^2 + \widetilde{m^2(t)}}\,100\%$$

For the special case of tone modulation,

$$m(t) = \mu A\cos\omega_m t \qquad\text{and}\qquad \widetilde{m^2(t)} = \frac{(\mu A)^2}{2}$$

Hence

$$\eta = \frac{\mu^2}{2 + \mu^2}\,100\%$$

with the condition that $0 \leq \mu \leq 1$. It can be seen that η increases monotonically with μ, and $\eta_{\max}$ occurs at $\mu = 1$, for which

$$\eta_{\max} = 33\%$$

Thus, for tone modulation, under the best conditions ($\mu = 1$), only one-third of the transmitted power is used for carrying messages. For practical signals, the efficiency is even worse—on the order of 25% or lower— in comparison to the DSB-SC case. The best condition implies $\mu = 1$. Smaller values of μ degrade efficiency further. For this reason, volume compression and peak limiting are commonly used in AM to ensure that full modulation ($\mu = 1$) is maintained most of the time.

Example 4.5 Determine η and the percentage of the total power carried by the sidebands of the AM wave for tone modulation when **(a)** $\mu = 0.5$ and **(b)** $\mu = 0.3$.

For $\mu = 0.5$,

$$\eta = \frac{\mu^2}{2+\mu^2}100\% = \frac{(0.5)^2}{2+(0.5)^2}100\% = 11.11\%$$

Hence, only about 11% of the total power is in the sidebands. For $\mu = 0.3$,

$$\eta = \frac{(0.3)^2}{2+(0.3)^2}100\% = 4.3\%$$

Hence, only 4.3% of the total power is the useful information power (in sidebands).

Generation of AM Signals

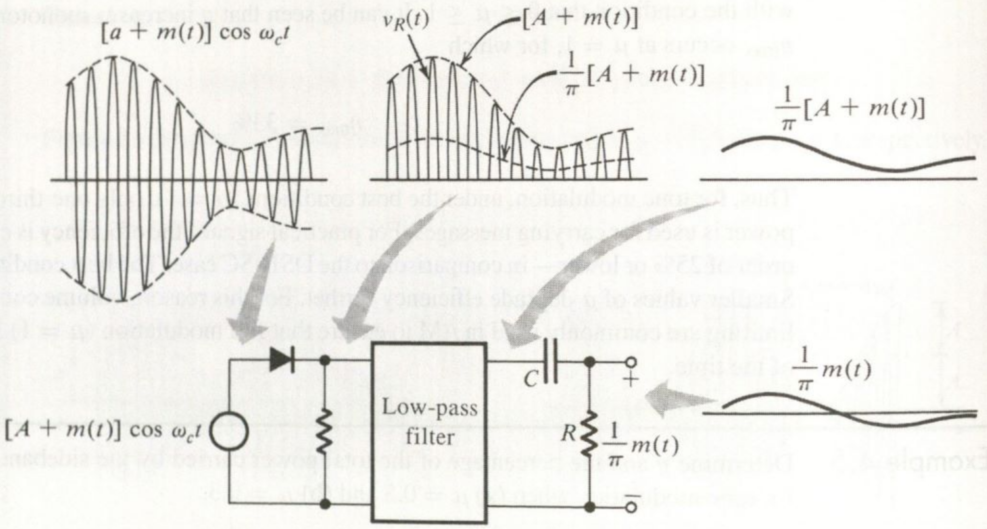
In principle, the generation of AM signals is identical to any DSB-SC modulation discussed in Sec. 4.2 except that an additional carrier component $A \cos \omega_c t$ needs to be added to the DSB-SC signal.

Demodulation of AM Signals

Like DSB-SC signals, the AM signal can be demodulated coherently by a locally generated carrier. Coherent, or synchronous, demodulation of AM defeats the purpose of AM, however, because it does not take advantage of the additional carrier component $A \cos \omega_c t$. As we have seen earlier, in the case of $\mu \leq 1$, the envelope of AM signal follows the message signal $m(t)$. Hence, we shall consider two noncoherent methods of AM demodulation under the condition of $0 < \mu \leq 1$: (1) rectifier detection and (2) envelope detection.

Rectifier: If an AM signal is applied to a diode and a resistor circuit (Fig. 4.10), the negative part of the AM wave will be removed. The output across the resistor is a half-wave rectified version of the AM signal. Visually, the diode acts like a pair of scissors by cutting off any negative half-cycle of the modulated sinusoid. In essence, at the rectifier output, the AM

Figure 4.10
Rectifier detector for AM.


$[a + m(t)] \cos \omega_c t$
$v_R(t)$
$[A + m(t)]$
$\frac{1}{\pi}[A + m(t)]$
$\frac{1}{\pi}[A + m(t)]$
$[A + m(t)] \cos \omega_c t$
Low-pass filter
C
$+$
$\frac{1}{\pi} m(t)$
R
$\frac{1}{\pi} m(t)$

signal is multiplied by $w(t)$. Hence, the half-wave rectified output $v_R(t)$ is

$$v_R(t) = \{[A + m(t)] \cos \omega_c t\} \, w(t) \tag{4.12}$$

$$= [A + m(t)] \cos \omega_c t \left[\frac{1}{2} + \frac{2}{\pi} \left(\cos \omega_c t - \frac{1}{3} \cos 3\omega_c t + \frac{1}{5} \cos 5\omega_c t - \cdots \right) \right] \tag{4.13}$$

$$= \frac{1}{\pi}[A + m(t)] + \text{other terms of higher frequencies} \tag{4.14}$$

When $v_R(t)$ is applied to a low-pass filter of cutoff B Hz, the output is $[A + m(t)]/\pi$, and all the other terms of frequencies higher than B Hz are suppressed. The dc term A/π may be blocked by a capacitor (Fig. 4.10) to give the desired output $m(t)/\pi$. The output can be doubled by using a full-wave rectifier.

It is interesting to note that because of the multiplication with $w(t)$, rectifier detection is in effect synchronous detection performed without using a local carrier. The high carrier content in AM ensures that its zero crossings are periodic and the information about frequency and phase of the carrier at the transmitter is built in to the AM signal itself.

Envelope Detector: In an envelope detector, the output of the detector follows the envelope of the modulated signal. The simple circuit shown in Fig. 4.11a functions as an envelope detector. On the positive cycle of the input signal, the input grows and may exceed the charged voltage on the capacity $v_C(t)$, turning on the diode and allowing the capacitor C to charge up to the peak voltage of the input signal cycle. As the input signal falls below this peak value, it falls quickly below the capacitor voltage (which is very nearly the peak voltage), thus causing the diode to open. The capacitor now discharges through the resistor R at a slow rate (with a time constant RC). During the next positive cycle, the same drama repeats. As the input signal rises above the capacitor voltage, the diode conducts again. The capacitor again charges to the peak value of this (new) cycle. The capacitor discharges slowly during the cutoff period.

During each positive cycle, the capacitor charges up to the peak voltage of the input signal and then decays slowly until the next positive cycle, as shown in Fig. 4.11b. Thus, the output voltage $v_C(t)$, closely follows the (rising) envelope of the input AM signal. Equally important, the slow capacity discharge via the resistor R allows the capacity voltage to follow

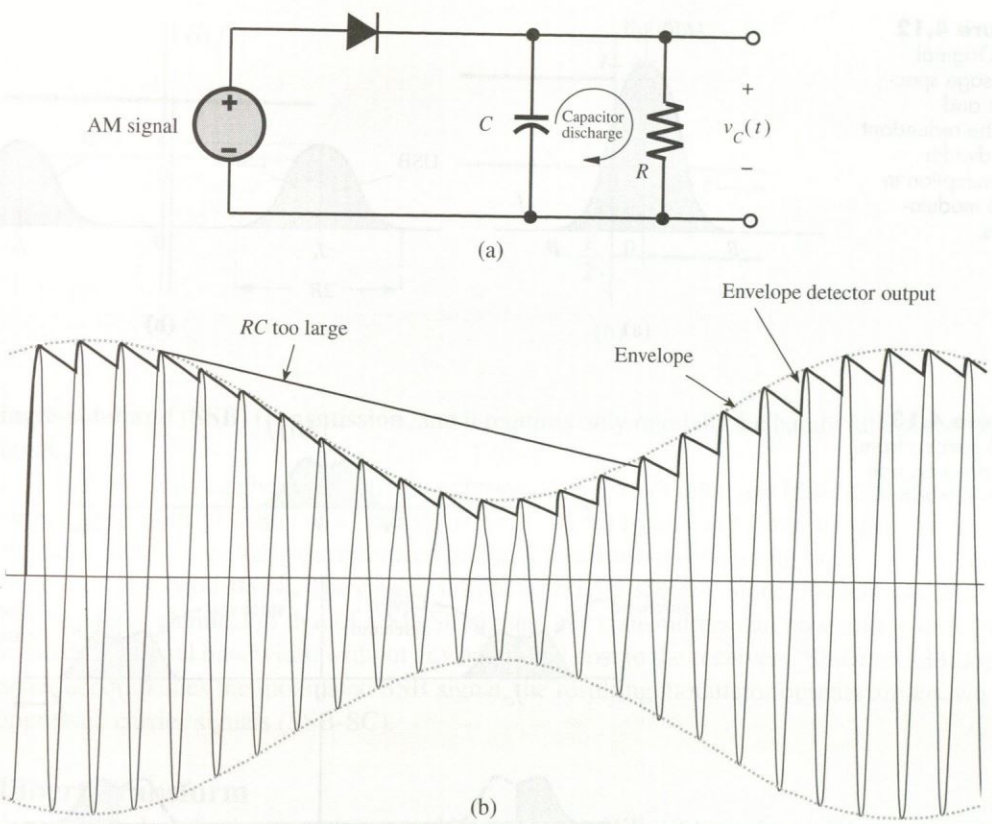


Figure 4.11 Envelope detector for AM.

a declining envelope. Capacitor discharge between positive peaks causes a ripple signal of frequency ω_c in the output. This ripple can be reduced by choosing a larger time constant RC so that the capacitor discharges very little between the positive peaks ($RC \gg 1/\omega_c$). If RC were made too large, however, it would be impossible for the capacitor voltage to follow a fast declining envelope (Fig. 4.11b). Because the maximum rate of AM envelope decline is dominated by the bandwidth B of the message signal $m(t)$, the design criterion of RC should be

$$1/\omega_c \ll RC < 1/(2\pi B) \qquad \text{or} \qquad 2\pi B < \frac{1}{RC} \ll \omega_c$$

The envelope detector output is $v_C(t) = A + m(t)$ with a ripple of frequency ω_c. The dc term A can be blocked out by a capacitor or a simple RC high-pass filter. The ripple may be reduced further by another (low-pass) RC filter.

4.4 BANDWIDTH-EFFICIENT AMPLITUDE MODULATIONS

As seen from Fig. 4.12, the DSB spectrum (including suppressed carrier and AM) has two sidebands: the upper sideband (USB) and the lower sideband (LSB), both containing complete information about the baseband signal $m(t)$. As a result, for a baseband signal $m(t)$ with bandwidth B Hz, DSB modulations require twice the radio-frequency bandwidth to transmit.

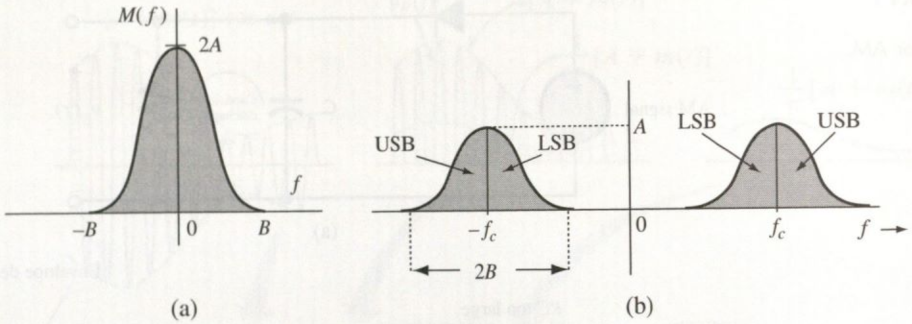

M(f)
2A
f
−B
0
B
(a)
USB
LSB
A
LSB
USB
−f_c
0
f_c
f →
2B
(b)

Figure 4.12 (a) Original message spectrum and (b) the redundant bandwidth consumption in DSB modulations.

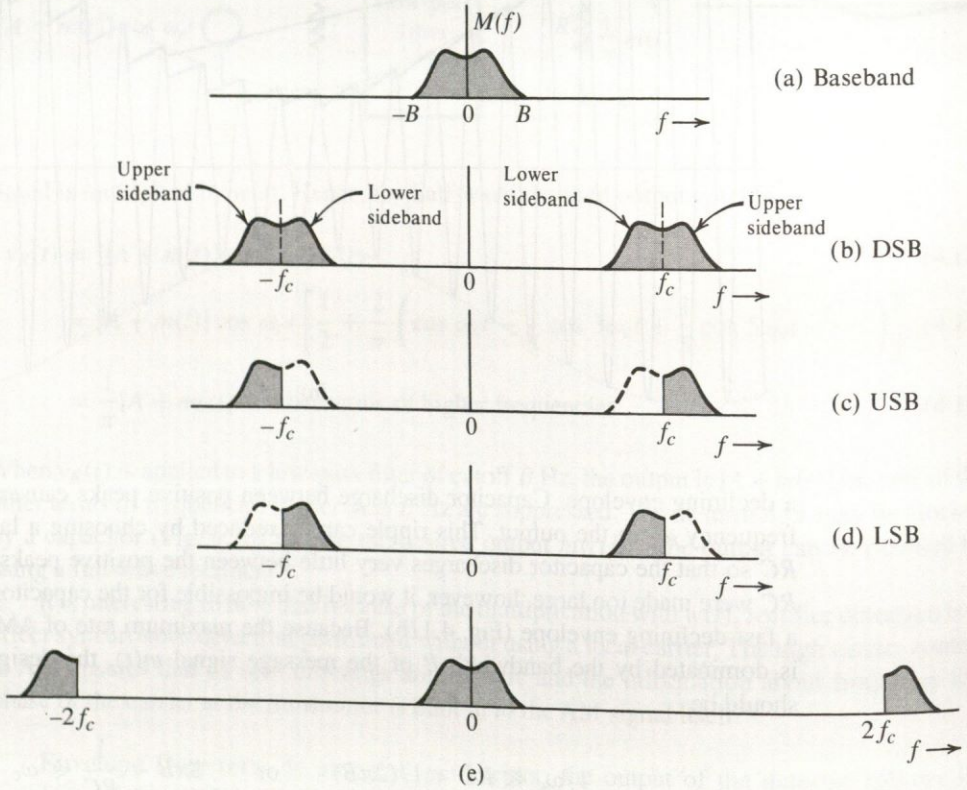

M(f)
(a) Baseband
−B
0
B
f →
Upper sideband
Lower sideband
Lower sideband
Upper sideband
−f_c
0
f_c
f →
(b) DSB
(c) USB
(d) LSB
−2f_c
2f_c
(e)

Figure 4.13 SSB spectra from suppressing one DSB sideband.

To improve the spectral efficiency of amplitude modulation, there exist two basic schemes to either utilize or remove the 100% spectral redundancy:

- Single-sideband (SSB) modulation, which removes either the LSB or the USB so that for one message signal $m(t)$, there is only a bandwidth of B Hz.
- Quadrature amplitude modulation (QAM), which utilizes spectral redundancy by sending two messages over the same bandwidth of $2B$ Hz.

4.4.1 Amplitude Modulation: Single Sideband (SSB)

As shown in Fig. 4.13, either the LSB or the USB can be suppressed from the DSB signal via bandpass filtering. Such a scheme, in which only one sideband is transmitted, is known as

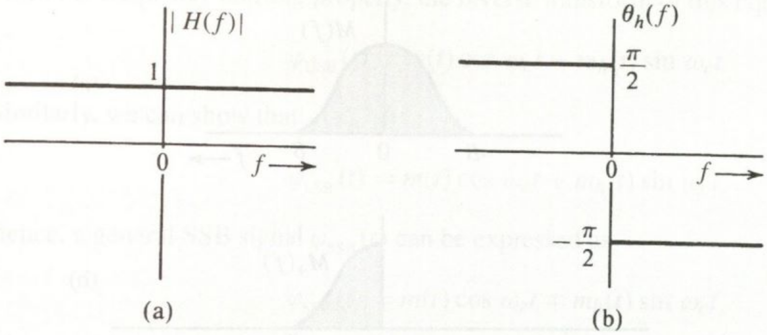


Figure 4.14 Transfer function of an ideal $\pi/2$ phase shifter (Hilbert transformer).

single-sideband (SSB) transmission, and it requires only one-half the bandwidth of the DSB signal.

An SSB signal can be coherently (synchronously) demodulated just like DSB-SC signals. For example, multiplication of a USB signal (Fig. 4.13c) by cos $\omega_c t$ shifts its spectrum to the left and right by f_c, yielding the spectrum in Fig. 4.13e. Low-pass filtering of this signal yields the desired baseband signal. The case is similar with LSB signals. Since the demodulation of SSB signals is identical to that of DSB-SC signals, the transmitters can now utilize only half the DSB-SC signal bandwidth without any additional cost to the receivers. Since no additional carrier accompanies the modulated SSB signal, the resulting modulator outputs are known as suppressed carrier signals (SSB-SC).

Hilbert Transform

We now introduce for later use a new tool known as the **Hilbert transform**. We use $x_h(t)$ and $\mathcal{H}\{x(t)\}$ to denote the Hilbert transform of a signal $x(t)$

$$x_h(t) = \mathcal{H}\{x(t)\} = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{x(\alpha)}{t-\alpha}\, d\alpha \tag{4.15}$$

Observe that the right-hand side of Eq. (4.15) has the form of a convolution

$$x(t) * \frac{1}{\pi t}$$

Now, application of the duality property to pair 12 of Table 3.1 yields $1/\pi t \Longleftrightarrow -j\,\mathrm{sgn}\,(f)$. Hence, application of the time convolution property to Eq. (4.15) convolution yields

$$X_h(f) = -jX(f)\ \mathrm{sgn}\ (f) \tag{4.16}$$

From Eq. (4.16), it follows that if $m(t)$ is passed through a transfer function $H(f) = -j\ \mathrm{sgn}\ (f)$, then the output is $m_h(t)$, the Hilbert transform of $m(t)$. Because

$$H(f) = -j\ \mathrm{sgn}\ (f) \tag{4.17}$$

$$= \begin{cases} -j = 1 \cdot e^{-j\pi/2} & f > 0 \\ j = 1 \cdot e^{j\pi/2} & f < 0 \end{cases} \tag{4.18}$$

it follows that $|H(f)| = 1$ and that $\theta_h(f) = -\pi/2$ for $f > 0$ and $\pi/2$ for $f < 0$, as shown in Fig. 4.14. Thus, if we change the phase of every component of $m(t)$ by $\pi/2$ (without changing any component's amplitude), the resulting signal is $m_h(t)$, the Hilbert transform of

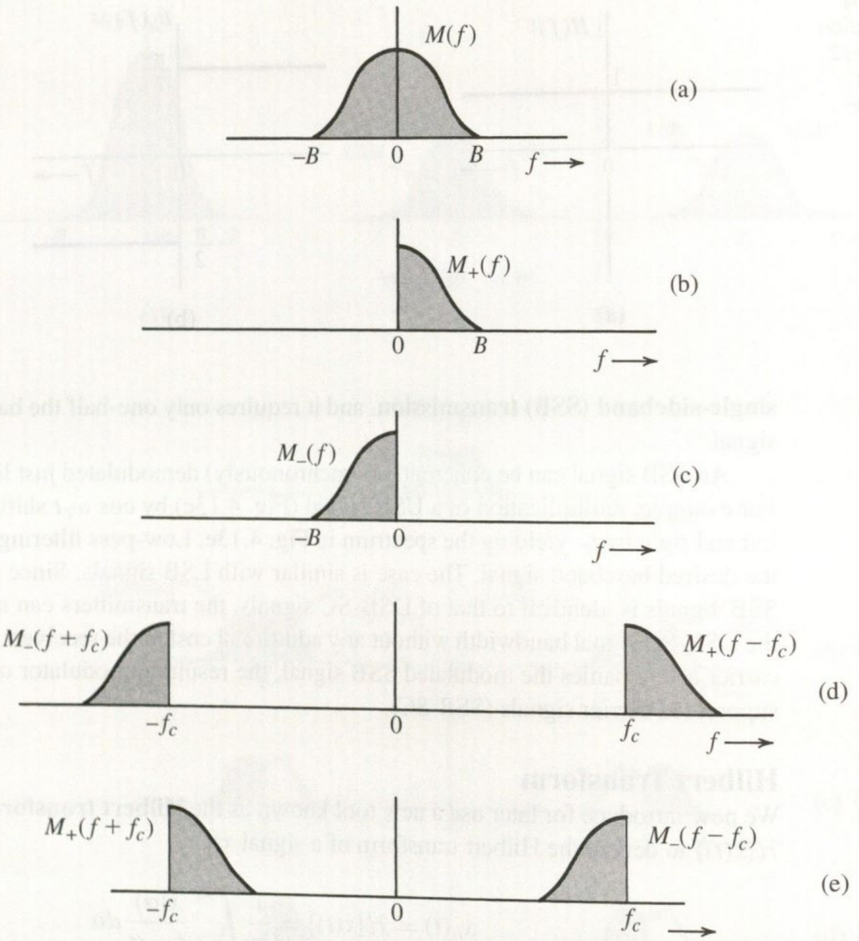


Figure 4.15 Expressing SSB spectra in terms of $M_+(f)$ and $M_-(f)$.

$m(t)$. Therefore, a Hilbert transformer is an ideal phase shifter that shifts the phase of every positive spectral component by $-\pi/2$.

Time Domain Representation of SSB Signals

Because the building blocks of an SSB signal are the sidebands, we shall first obtain a time domain expression for each sideband.

Figure 4.15a shows the message spectrum $M(f)$. Figure 4.15b shows its right half $M_+(f)$ and Fig. 4.15c shows its left half $M_-(f)$. From Fig. 4.15b and c, we observe that

$$M_+(f) = M(f) \cdot u(f) = M(f)\frac{1}{2}[1 + \text{sgn}(f)] = \frac{1}{2}[M(f) + jM_h(f)] \tag{4.19a}$$

$$M_-(f) = M(f)u(-f) = M(f)\frac{1}{2}[1 - \text{sgn}(f)] = \frac{1}{2}[M(f) - jM_h(f)] \tag{4.19b}$$

We can now express the SSB signal in terms of $m(t)$ and $m_h(t)$. From Fig. 4.15d it is clear that the USB spectrum $\Phi_{\text{USB}}(f)$ can be expressed as

$$\begin{aligned}\Phi_{\text{USB}}(f) &= M_+(f - f_c) + M_-(f + f_c)\\ &= \frac{1}{2}[M(f - f_c) + M(f + f_c)] - \frac{1}{2j}[M_h(f - f_c) - M_h(f + f_c)]\end{aligned}$$

From the frequency-shifting property, the inverse transform of this equation yields

$$\varphi_{\text{USB}}(t) = m(t)\cos\omega_c t - m_h(t)\sin\omega_c t \tag{4.20a}$$

Similarly, we can show that

$$\varphi_{\text{LSB}}(t) = m(t)\cos\omega_c t + m_h(t)\sin\omega_c t \tag{4.20b}$$

Hence, a general SSB signal $\varphi_{\text{SSB}}(t)$ can be expressed as

$$\varphi_{\text{SSB}}(t) = m(t)\cos\omega_c t \mp m_h(t)\sin\omega_c t \tag{4.20c}$$

where the minus sign applies to USB and the plus sign applies to LSB.

Given the time-domain expression of SSB-SC signals, we can now confirm analytically (instead of graphically) that SSB-SC signals can be coherently demodulated:

$$\begin{aligned}
\varphi_{\text{SSB}}(t)\cos\omega_c t &= [m(t)\cos\omega_c t \mp m_h(t)\sin\omega_c t]\,2\cos\omega_c t\\
&= m(t)[1+\cos 2\omega_c t] \mp m_h(t)\sin 2\omega_c t\\
&= m(t) + \underbrace{[m(t)\cos 2\omega_c t \mp m_h(t)\sin 2\omega_c t]}_{\text{SSB-SC signal with carrier frequency } 2\omega_c}
\end{aligned}$$

Thus, the product $\varphi_{\text{SSB}}(t)\cdot 2\cos\omega_c t$ yields the baseband signal and another SSB signal with a carrier frequency $2\omega_c$. The spectrum in Fig. 4.13e shows precisely this result. A low-pass filter will suppress the unwanted SSB terms, giving the desired baseband signal $m(t)$. Hence, the demodulator is identical to the synchronous demodulator used for DSB-SC. Thus, any one of the synchronous DSB-SC demodulators discussed earlier in Sec. 4.2 can be used to demodulate an SSB-SC signal.

Example 4.6 Tone Modulation: SSB

Find $\varphi_{\text{SSB}}(t)$ for the simple case of a tone modulation, that is, a modulating signal that is a sinusoid $m(t) = \cos\omega_m t$. Also demonstrate the coherent demodulation of this SSB signal.

Recall that the Hilbert transform delays the phase of each spectral component by $\pi/2$. In the present case, there is only one spectral component, of frequency ω_m. Delaying the phase of $m(t)$ by $\pi/2$ yields

$$m_h(t) = \cos\left(\omega_m t - \frac{\pi}{2}\right) = \sin\omega_m t$$

Hence, from Eq. (4.20c),

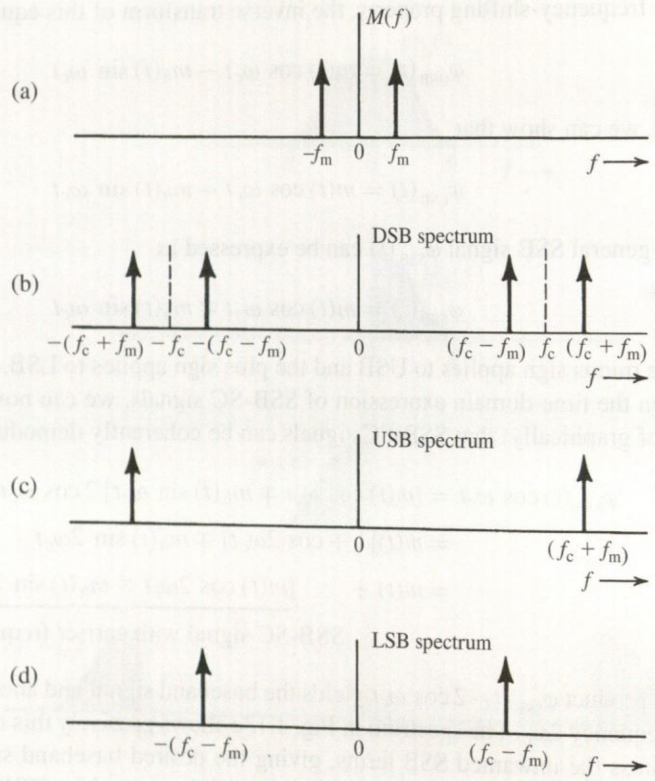
$$\begin{aligned}
\varphi_{\text{SSB}}(t) &= \cos\omega_m t\cos\omega_c t \mp \sin\omega_m t\sin\omega_c t\\
&= \cos(\omega_c \pm \omega_m)t
\end{aligned}$$

Thus,

$$\varphi_{\text{USB}}(t) = \cos(\omega_c + \omega_m)t \quad \text{and} \quad \varphi_{\text{LSB}}(t) = \cos(\omega_c - \omega_m)t$$

To verify these results, consider the spectrum of $m(t)$ (Fig. 4.16a) and its DSB-SC (Fig. 4.16b), USB (Fig. 4.16c), and LSB (Fig. 4.16d) spectra. It is evident that the spectra in Fig. 4.16c and d do indeed correspond to the $\varphi_{\text{USB}}(t)$ and $\varphi_{\text{LSB}}(t)$ derived earlier.

Figure 4.16 SSB spectra for tone modulation.



Finally, the coherent demodulation of the SSB tone modulation can be achieved by

$$\begin{aligned}\varphi_{\text{SSB}}(t) 2\cos\omega_c t &= 2\cos(\omega_c \pm \omega_m)t \cos\omega_c t \\ &= \cos\omega_m t + \cos(\omega_c + \omega_m)t\end{aligned}$$

which can be sent to a low-pass filter to retrieve the message tone $\cos\omega_m t$.

SSB Modulation Systems

Three methods are commonly used to generate SSB signals: phase shifting, selective filtering, and the Weaver method. None of these modulation methods are precise, and each one generally requires that the baseband signal spectrum have little power near the origin.

The **phase shift method** directly uses Eq. (4.20) as its basis. In Fig. 4.17, which shows its implementation, $-\pi/2$ designates a $\mp\pi/2$ phase shifter, which delays the phase of every positive spectral component by $\pi/2$. Hence, it is a Hilbert transformer. Note that an ideal Hilbert phase shifter is unrealizable. This is because of the requirement for an abrupt phase change of π at zero frequency. When the message $m(t)$ has a dc null and very little low-frequency content, the practical approximation of this ideal phase shifter has almost no real effect and does not affect the accuracy of SSB modulation.

In the **selective-filtering method**, the most commonly used method of generating SSB signals, a DSB-SC signal is passed through a sharp cutoff filter to eliminate the undesired sideband. To obtain the USB, the filter should pass all components above frequency f_c unattenuated and completely suppress all components below frequency f_c. Such an operation requires an

Figure 4.17 Using the phase-shift method to generate SSB.

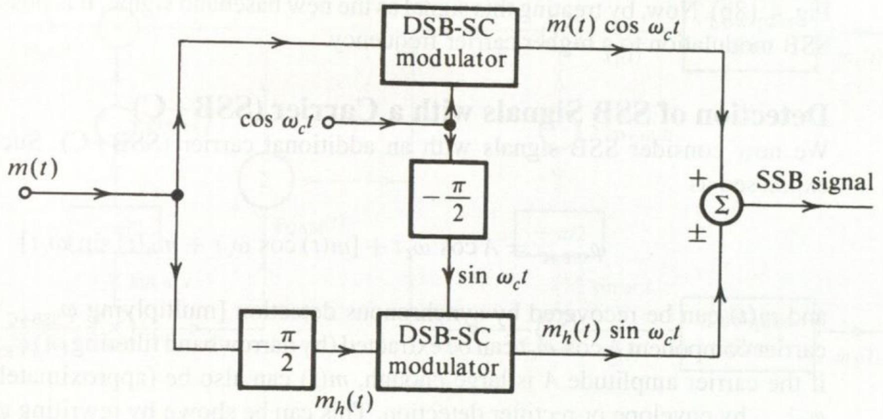
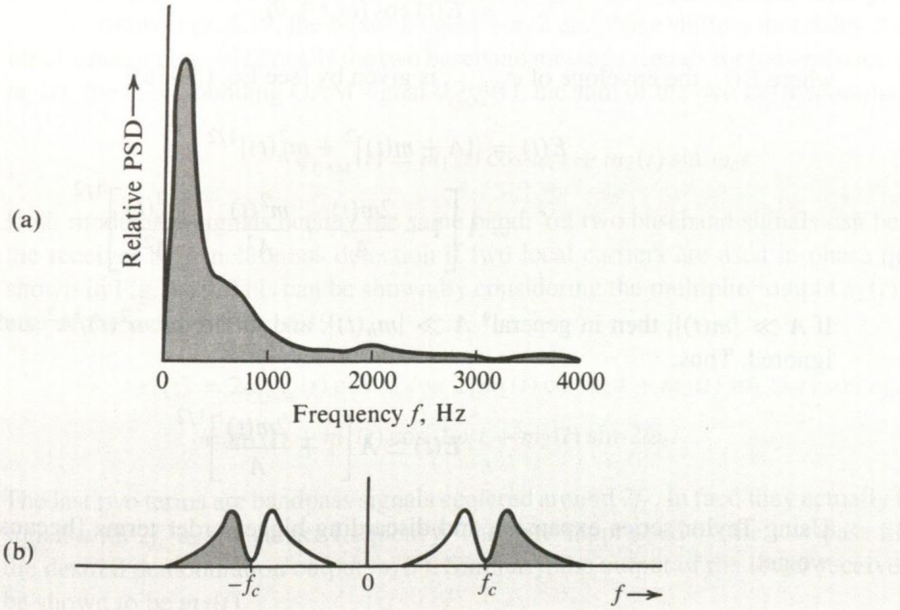


Figure 4.18 (a) Relative power spectrum of speech signal and (b) the corresponding USB spectrum.



ideal filter, which is unrealizable. It can, however, be approximated closely if there is some separation between the passband and the stopband. Fortunately, the voice signal provides this condition, because its spectrum shows little power content at the origin (Fig. 4.18a). In addition, articulation tests have shown that for speech signals, frequency components below 300 Hz are not important. In other words, we may suppress all speech components below 300 Hz (and above 3500 Hz) without affecting intelligibility appreciably. Thus, filtering of the unwanted sideband becomes relatively easy for speech signals because we have a 600 Hz transition region around the cutoff frequency f_c. To minimize adjacent channel interference, the undesired sideband should be attenuated at least 40 dB.

For very high carrier frequency f_c, the ratio of the gap band (600 Hz) to the carrier frequency may be too small, and, thus, a transition of 40 dB in amplitude over 600 Hz may be difficult. In such a case, a third method, known as **Weaver's method**,[1] utilizes two stages of SSB amplitude modulation. First, the modulation is carried out using a smaller carrier frequency f_{c_1}. The resulting SSB signal effectively widens the gap to $2f_{c_1}$ (shaded spectrum in

Fig. 4.18b). Now, by treating this signal as the new baseband signal, it is possible to accomplish SSB modulation to a higher carrier frequency.

Detection of SSB Signals with a Carrier (SSB+C)

We now consider SSB signals with an additional carrier (SSB+C). Such a signal can be expressed as

$$\varphi_{\text{SSB+C}} = A\cos\omega_c t + [m(t)\cos\omega_c t + m_h(t)\sin\omega_c t]$$

and $m(t)$ can be recovered by synchronous detection [multiplying $\varphi_{\text{SSB+C}}$ by $\cos\omega_c t$] if the carrier component $A\cos\omega_c t$ can be extracted (by narrowband filtering of) $\varphi_{\text{SSB+C}}$. Alternatively, if the carrier amplitude A is large enough, $m(t)$ can also be (approximately) recovered from $\varphi_{\text{SSB+C}}$ by envelope or rectifier detection. This can be shown by rewriting $\varphi_{\text{SSB+C}}$ as

$$\begin{aligned}\varphi_{\text{SSB+C}} &= [A + m(t)]\cos\omega_c t + m_h(t)\sin\omega_c t\\ &= E(t)\cos(\omega_c t + \theta)\end{aligned} \tag{4.21}$$

where $E(t)$, the envelope of $\varphi_{\text{SSB+C}}$, is given by [see Eq. (3.41a)]

$$\begin{aligned}E(t) &= \{[A + m(t)]^2 + m_h^2(t)\}^{1/2}\\ &= A\left[1 + \frac{2m(t)}{A} + \frac{m^2(t)}{A^2} + \frac{m_h^2(t)}{A^2}\right]^{1/2}\end{aligned}$$

If $A \gg |m(t)|$, then in general* $A \gg |m_h(t)|$, and the terms $m^2(t)/A^2$ and $m_h^2(t)/A^2$ can be ignored. Thus,

$$E(t) \simeq A\left[1 + \frac{2m(t)}{A}\right]^{1/2}$$

Using Taylor series expansion and discarding higher order terms [because $|m(t)/A| \ll 1$], we get

$$\begin{aligned}E(t) &\simeq A\left[1 + \frac{m(t)}{A}\right]\\ &= A + m(t)\end{aligned}$$

It is evident that for a large carrier, the SSB+C can be demodulated by an envelope detector.

In AM, envelope detection requires the condition $A \geq |m(t)|$, whereas for SSB+C, the condition is $A \gg |m(t)|$. Hence, in SSB case, the required carrier amplitude is much larger than that in AM, and, consequently, the efficiency of SSB+C is pathetically low.

4.4.2 Quadrature Amplitude Modulation (QAM)

Because SSB-SC signals are difficult to generate accurately, quadrature amplitude modulation (QAM) offers an attractive alternative to SSB-SC. QAM can be exactly generated without

---

* This may not be true for all t, but it is true for most t.

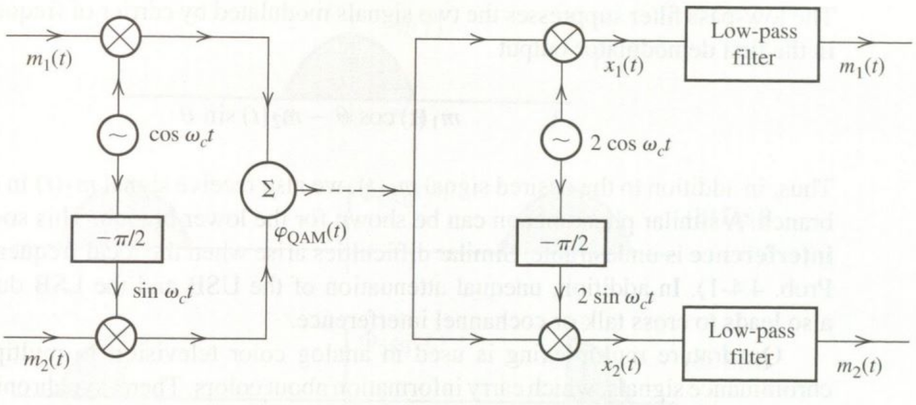


Figure 4.19 Quadrature amplitude modulation.

requiring sharp cutoff bandpass filters. QAM operates by transmitting two DSB signals via carriers of the same frequency but in phase quadrature, as shown in Fig. 4.19. This scheme is known as **quadrature amplitude modulation (QAM)**, or **quadrature multiplexing**.

As shown Fig. 4.19, the boxes labeled $-\pi/2$ are phase shifters that delay the phase of an input sinusoid by $-\pi/2$ rad. If the two baseband message signals for transmission are $m_1(t)$ and $m_2(t)$, the corresponding QAM signal $\varphi_{\text{QAM}}(t)$, the sum of the two DSB-modulated signals, is

$$\varphi_{\text{QAM}}(t) = m_1(t) \cos \omega_c t + m_2(t) \sin \omega_c t$$

Both modulated signals occupy the same band. Yet two baseband signals can be separated at the receiver by synchronous detection if two local carriers are used in phase quadrature, as shown in Fig. 4.19. This can be shown by considering the multiplier output $x_1(t)$ of the upper arm of the receiver (Fig. 4.19):

$$\begin{aligned} x_1(t) &= 2\varphi_{\text{QAM}}(t) \cos \omega_c t = 2[m_1(t) \cos \omega_c t + m_2(t) \sin \omega_c t] \cos \omega_c t \\ &= m_1(t) + m_1(t) \cos 2\omega_c t + m_2(t) \sin 2\omega_c t \end{aligned} \tag{4.22a}$$

The last two terms are bandpass signals centered around $2f_c$. In fact, they actually form a QAM signal with $2f_c$ as the carrier frequency. They are suppressed by the low-pass filter, yielding the desired demodulation output $m_1(t)$. Similarly, the output of the lower receiver branch can be shown to be $m_2(t)$.

$$\begin{aligned} x_2(t) &= 2\varphi_{\text{QAM}}(t) \sin \omega_c t = 2[m_1(t) \cos \omega_c t + m_2(t) \sin \omega_c t] \sin \omega_c t \\ &= m_2(t) - m_2(t) \cos 2\omega_c t + m_1(t) \sin 2\omega_c t \end{aligned} \tag{4.22b}$$

Thus, two baseband signals, each of bandwidth B Hz, can be transmitted simultaneously over a bandwidth $2B$ by using DSB transmission and quadrature multiplexing. The upper channel is also known as the **in-phase** (I) channel and the lower channel is the **quadrature** (Q) channel. Both signals $m_1(t)$ and $m_2(t)$ can be separately demodulated.

Note, however, that QAM demodulation must be totally synchronous. An error in the phase or the frequency of the carrier at the demodulator in QAM will result in loss and interference between the two channels. To show this let the carrier at the demodulator be $2 \cos (\omega_c t + \theta)$. In this case,

$$\begin{aligned} x_1(t) &= 2[m_1(t) \cos \omega_c t + m_2(t) \sin \omega_c t] \cos (\omega_c t + \theta) \\ &= m_1(t) \cos \theta - m_2(t) \sin \theta + m_1(t) \cos (2\omega_c t + \theta) + m_2(t) \sin (2\omega_c t + \theta) \end{aligned}$$

The low-pass filter suppresses the two signals modulated by carrier of frequency $2f_c$, resulting in the first demodulator output

$$m_1(t)\cos\theta - m_2(t)\sin\theta$$

Thus, in addition to the desired signal $m_1(t)$, we also receive signal $m_2(t)$ in the upper receiver branch. A similar phenomenon can be shown for the lower branch. This so-called **cochannel interference** is undesirable. Similar difficulties arise when the local frequency is in error (see Prob. 4.4-1). In addition, unequal attenuation of the USB and the LSB during transmission also leads to cross talk or cochannel interference.

Quadrature multiplexing is used in analog color television to multiplex the so-called chrominance signals, which carry information about colors. There, synchronization is achieved by periodic insertion of a short burst of carrier signal (called a **color burst** in the transmitted signal, as explained in Sec. 4.9). Digital satellite television transmission also applies QAM.

With respect to bandwidth requirement, SSB is similar to QAM but less exacting in terms of the carrier frequency and phase or the requirement of a distortionless transmission medium. However, SSB is difficult to generate if the baseband signal $m(t)$ has significant spectral content near the direct current (dc).

4.5 AMPLITUDE MODULATIONS: VESTIGIAL SIDEBAND (VSB)

As discussed earlier, the generation of exact SSB signals is rather difficult: generally, the message signal $m(t)$ must have a null around direct current. A phase shifter required in the phase shift method is unrealizable, or realizable only approximately. The generation of DSB signals is much simpler but requires twice the signal bandwidth. The **vestigial-sideband (VSB)** modulation system, also called asymmetric sideband, is a compromise between DSB and SSB. It inherits the advantages of DSB and SSB but avoids their disadvantages at a small cost. VSB signals are relatively easy to generate, and, at the same time, their bandwidth is only some what (typically 25%) greater than that of SSB signals.

In VSB, instead of rejecting one sideband completely (as in SSB), a gradual cutoff of one sideband, as shown in Fig. 4.20d, is accepted. The baseband signal can be recovered exactly by a synchronous detector in conjunction with an appropriate equalizer filter $H_o(f)$ at the receiver output (Fig. 4.21). If a large carrier is transmitted along with the VSB signal, the baseband signal can even be recovered by an envelope (or a rectifier) detector.

If the vestigial shaping filter that produces VSB from DSB is $H_i(f)$ (Fig. 4.21), then the resulting VSB signal spectrum is

$$\Phi_{\text{VSB}}(f) = [M(f+f_c) + M(f-f_c)]H_i(f) \tag{4.23}$$

This VSB shaping filter $H_i(f)$ allows the transmission of one sideband but suppresses the other sideband, not completely, but gradually. This makes it easy to realize such a filter, but the transmission bandwidth is now somewhat higher than that of the SSB (where the other sideband is suppressed completely). The bandwidth of the VSB signal is typically 25 to 33% higher than that of the SSB signals.

We require that $m(t)$ be recoverable from $\varphi_{\text{VSB}}(t)$ using synchronous demodulation at the receiver. This is done by multiplying the incoming VSB signal $\varphi_{\text{VSB}}(t)$ by $2\cos\omega_c t$. The

Figure 4.20 Spectra of the modulating signal and corresponding DSB, SSB, and VSB signals.

Figure 4.21 VSB modulator and demodulator.

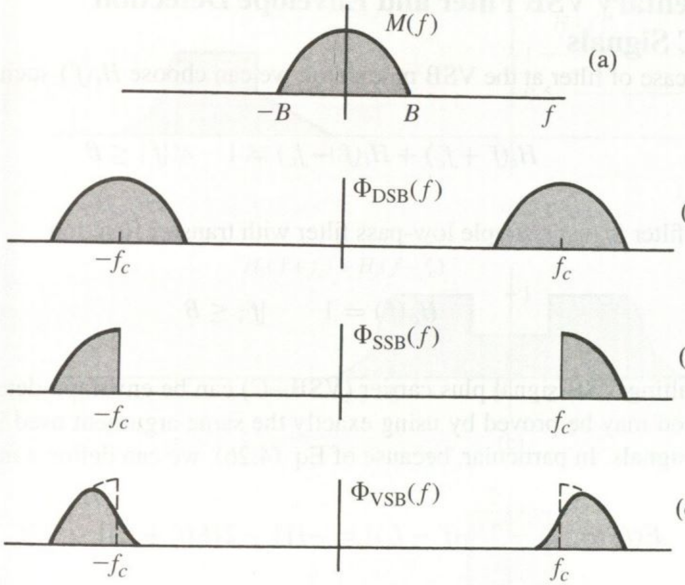


Figure 4.20 Spectra of the modulating signal and corresponding DSB, SSB, and VSB signals.

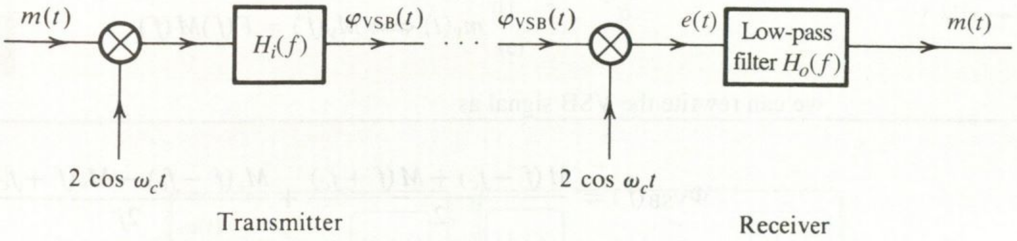


Figure 4.21 VSB modulator and demodulator.

product $e(t)$ is given by

$$e(t) = 2\varphi_{\text{VSB}}(t)\cos\omega_c t \Longleftrightarrow [\Phi_{\text{VSB}}(f+f_c) + \Phi_{\text{VSB}}(f-f_c)]$$

The signal $e(t)$ is further passed through the low-pass equalizer filter of transfer function $H_o(f)$. The output of the equalizer filter is required to be $m(t)$. Hence, the output signal spectrum is given by

$$M(f) = [\Phi_{\text{VSB}}(f+f_c) + \Phi_{\text{VSB}}(f-f_c)]H_o(f)$$

Substituting Eq. (4.23) into this equation and eliminating the spectra at $\pm 2f_c$ [suppressed by a low-pass filter $H_o(f)$], we obtain

$$M(f) = M(f)[H_i(f+f_c) + H_i(f-f_c)]H_o(f) \tag{4.24}$$

Hence

$$H_o(f) = \frac{1}{H_i(f+f_c) + H_i(f-f_c)} \qquad |f| \leq B \tag{4.25}$$

Note that because $H_i(f)$ is a bandpass filter, the terms $H_i(f \pm f_c)$ contain low-pass components.

Complementary VSB Filter and Envelope Detection of VSB+C Signals

As a special case of filter at the VSB modulator, we can choose $H_i(f)$ such that

$$H_i(f+f_c) + H_i(f-f_c) = 1 \qquad |f| \le B \tag{4.26}$$

Then output filter is just a simple low-pass filter with transfer function

$$H_o(f) = 1 \qquad |f| \le B$$

The resulting VSB signal plus carrier (VSB+C) can be envelope-detected. This demodulation method may be proved by using exactly the same argument used in proving the case for SSB+C signals. In particular, because of Eq. (4.26), we can define a new low-pass filter

$$F(f) = j\,[1 - 2H_i(f-f_c)] = -j\,[1 - 2H_i(f+f_c)] \qquad |f| \le B$$

Upon defining a new (complex) low-pass signal as

$$m_v(t) \Longleftrightarrow M_v(f) = F(f)M(f)$$

we can rewrite the VSB signal as

$$\Phi_{\text{VSB}}(f) = \frac{M(f-f_c) + M(f+f_c)}{2} + \frac{M_v(f-f_c) - M_v(f+f_c)}{2j} \tag{4.27a}$$

$$\Longleftrightarrow$$

$$\varphi_{\text{VSB}}(t) = m(t)\cos 2\pi f_c t + m_v(t)\sin 2\pi f_c t \tag{4.27b}$$

Clearly, both the SSB and the VSB modulated signals have the same form, with $m_h(t)$ in SSB replaced by a low-pass signal $m_v(t)$ in VSB. Applying the same analysis from the SSB+C envelope detection, a large carrier addition to $\varphi_{\text{VSB}}(t)$ would allow the envelope detection of VSB+C.

We have shown that SSB+C requires a much larger carrier than DSB+C (AM) for envelope detection. Because VSB+C is an in-between case, the added carrier required in VSB is larger than that in AM, but smaller than that in SSB+C.

Example 4.7 The carrier frequency of a certain VSB signal is $f_c = 20$ kHz, and the baseband signal bandwidth is 6 kHz. The VSB shaping filter $H_i(f)$ at the input, which cuts off the lower sideband gradually over 2 kHz, is shown in Fig. 4.22a. Find the output filter $H_o(f)$ required for distortionless reception.

Figure 4.22b shows the low-pass segments of $H_i(f+f_c) + H_i(f-f_c)$. We are interested in this spectrum only over the baseband (the remaining undesired portion is suppressed by the output filter). This spectrum is 0.5 over the band of 0 to 2 kHz, and is 1 over 2 to 6 kHz, as shown in Fig. 4.22b. Figure 4.22c shows the desired output filter $H_o(f)$, which is the reciprocal of the spectrum in Fig. 4.22b [see Eq. (4.25)].

Figure 4.22
VSB modulator and receiver filters.

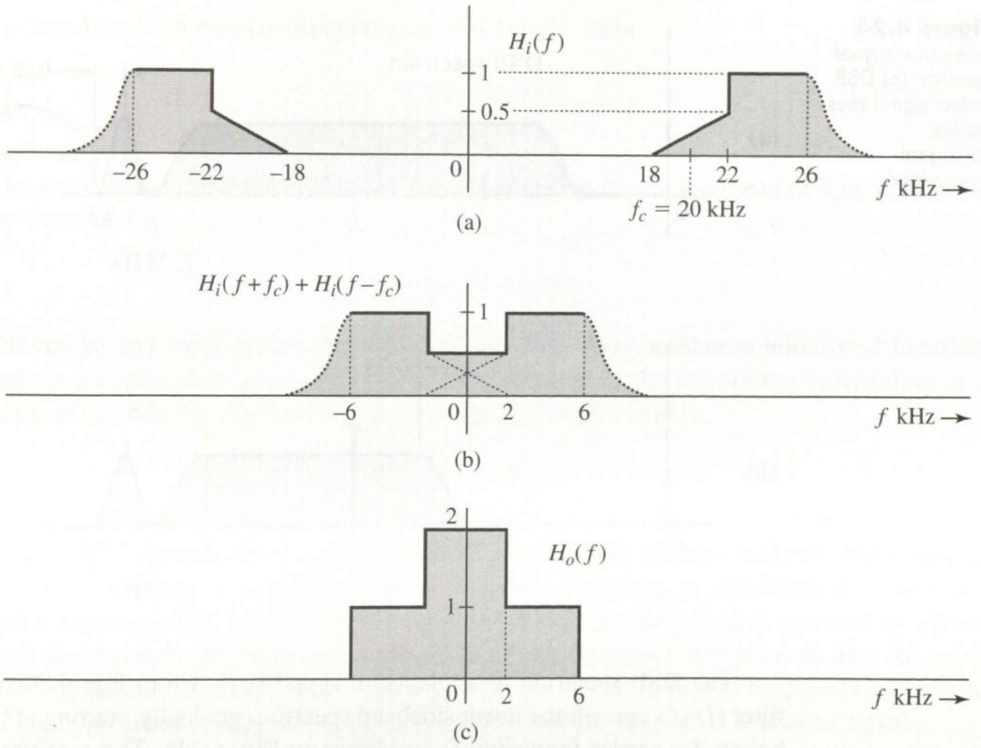
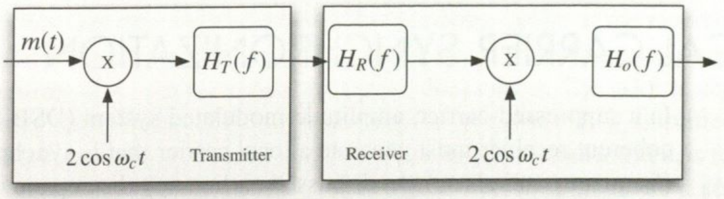


Figure 4.23
Transmitter filter $H_T(f)$, receiver front-end filter $H_R(f)$, and the receiver output low-pass filter $H_o(f)$ in VSB television systems.



Use of VSB in Broadcast Television

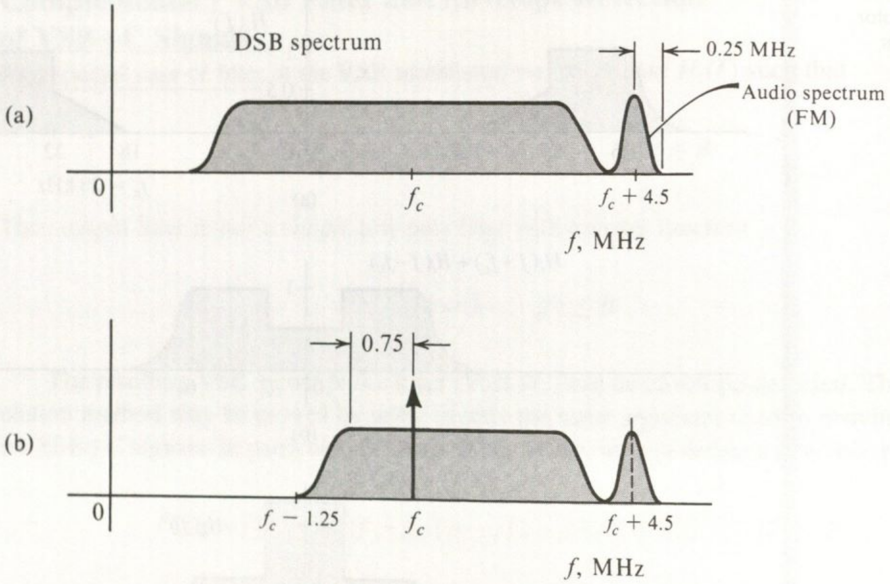
VSB is very attractive for television broadcast systems because it is a clever compromise between SSB and DSB. The baseband video signal of television occupies an enormous bandwidth of 4.5 MHz, and a DSB signal needs a bandwidth of 9 MHz. It would seem desirable to use SSB to conserve bandwidth. Unfortunately, this creates several problems. First, the baseband video signal has sizable power in the low-frequency region, and consequently it is difficult to suppress one sideband completely. Second, in designing a broadcast receiver, we prefer an envelope detector over a synchronous one to reduce the receiver cost. We saw earlier that SSB+C has a very low power efficiency; moreover, use of SSB would increase the receiver cost.

The spectral shaping of television VSB signals can be illustrated by Fig. 4.23. The vestigial spectrum is controlled by two filters: the transmitter RF filter $H_T(f)$ and the receiver RF filter $H_R(f)$. Jointly we have

$$H_i(f) = H_T(f)H_R(f)$$

Hence, the design of the receiver output filter $H_o(f)$ follows Eq. (4.25).

Figure 4.24
Television signal spectra: (a) DSB video signal plus audio;
(b) signal transmitted.



The DSB spectrum of a television signal is shown in Fig. 4.24a. The vestigial shaping filter $H_T(f)$ cuts off the lower sideband spectrum gradually, starting at 0.75 MHz to 1.25 MHz below the carrier frequency f_c, as shown in Fig. 4.24b. The receiver output filter $H_o(f)$ is designed according to Eq. (4.25). The resulting VSB spectrum bandwidth is 6 MHz. Compare this with the DSB bandwidth of 9 MHz and the SSB bandwidth of 4.5 MHz.

4.6 LOCAL CARRIER SYNCHRONIZATION

In a suppressed-carrier, amplitude-modulated system (DSB-SC, SSB-SC, and VSB-SC), the coherent receiver must generate a local carrier that is synchronous with the incoming carrier (frequency and phase). As discussed earlier, any discrepancy in the frequency or phase of the local carrier gives rise to distortion in the detector output.

Consider an SSB-SC case where a received signal is

$$m(t)\cos[(\omega_c + \Delta\omega)t + \delta] - m_h(t)\sin[(\omega_c + \Delta\omega)t + \delta]$$

because of propagation delay and Doppler frequency shift. The local carrier remains as $2\cos\omega_c t$. The product of the received signal and the local carrier is $e(t)$, given by

$$\begin{aligned} e(t) &= 2\cos\omega_c t\,[m(t)\cos(\omega_c t + \Delta\omega t + \delta) - m_h(t)\sin(\omega_c t + \Delta\omega t + \delta)] \\ &= m(t)\cos(\Delta\omega t + \delta) - m_h(t)\sin(\Delta\omega t + \delta) \\ &\quad + \underbrace{m(t)\cos[(2\omega_c + \Delta\omega)t + \delta] - m_h(t)\sin[(2\omega_c + \Delta\omega)t + \delta]}_{\text{bandpass SSB-SC signal around } 2\omega_c + \Delta\omega} \end{aligned} \tag{4.28}$$

The bandpass component is filtered out by the receiver low-pass filter, leaving the output $e_o(t)$ as

$$e_o(t) = m(t)\cos(\Delta\omega t + \delta) - m_h(t)\sin(\Delta\omega t + \delta) \tag{4.29}$$

If $\Delta\omega$ and δ are both zero (no frequency or phase error), then

$$e_o(t) = m(t)$$

as expected.

In practice, if the radio wave travels a distance of d meters at the speed of light c, then the phase delay is

$$\delta = -(\omega_c + \Delta\omega)d/c$$

which can be any value within the interval $[-\pi, +\pi]$. Two oscillators initially of identical frequency can also drift apart. Moreover, if the receiver or the transmitter is traveling at a velocity of v_e, then the maximum Doppler frequency shift would be

$$\Delta f_{\max} = \frac{v_e}{c} f_c$$

The velocity v_e depends on the actual vehicle (e.g. spacecraft, airplane, and car). For example, if the mobile velocity v_e is 108 km/h, then for carrier frequency at 100 MHz, the maximum Doppler frequency shift would be 10 Hz. Such a shift of every frequency component by a fixed amount $\Delta\omega$ destroys the harmonic relationship among frequency components. For $\Delta f = 10$ Hz, the components of frequencies 1000 and 2000 Hz will be shifted to frequencies 1010 and 2010 Hz. This destroys their harmonic relationship, and the quality of nonaudio signals.

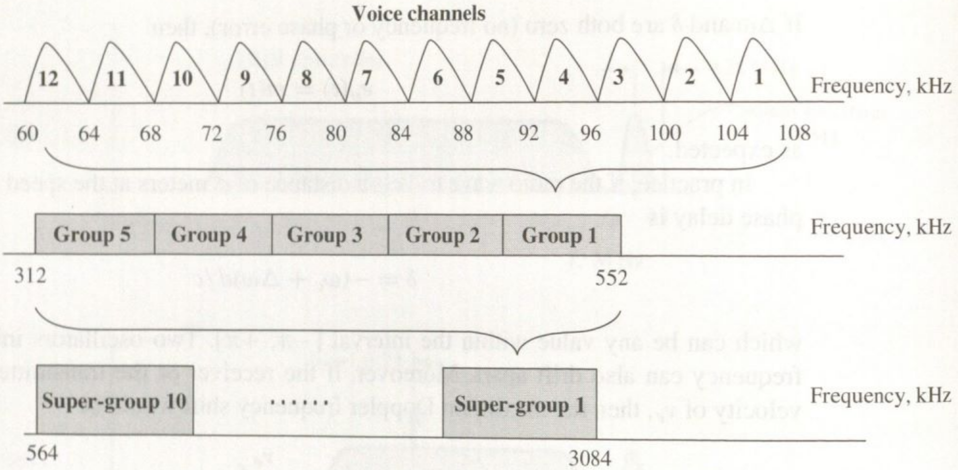
It is interesting to note that audio signals are highly redundant, and unless Δf is very large, such a change does not destroy the intelligibility of the output. For audio signals $\Delta f < 30$ Hz does not significantly affect the signal quality. A Δf exceeding 30 Hz results in a sound quality similar to that of Donald Duck. But intelligibility is not completely lost.

Generally, there are two ways to recover the incoming carrier at the receiver. One way is for the transmitter to transmit a pilot (sinusoid) signal that is either the exact carrier or directly related to the carrier (e.g., a pilot at half the carrier frequency). The pilot is separated at the receiver by a very narrowband filter tuned to the pilot frequency. It is amplified and used to synchronize the local oscillator. Another method, when no pilot is transmitted, is for the receiver to process the received signal by means of a nonlinear device to generate a separate carrier component to be extracted by means of narrow bandpass filters. Clearly, effective and narrow bandpass filters are very important to both methods. Moreover, the bandpass filter should also have the ability to adaptively adjust its center frequency to combat significant frequency drift or Doppler shift. Aside from some typical bandpass filter designs, the phase-locked loop (PLL), which plays an important role in carrier acquisition of various modulations, can be viewed as such a narrow and adaptive bandpass filter. The principles of PLL will be discussed later in this chapter.

4.7 FREQUENCY DIVISION MULTIPLEXING (FDM)

Signal multiplexing allows the transmission of several signals on the same channel. In Chapter 6, we shall discuss time division multiplexing (TDM), where several signals time-share the same channel. In FDM, several signals share the band of a channel. Each signal is modulated by a different carrier frequency. The various carriers are adequately separated to avoid overlap (or interference) between the spectra of various modulated signals. These carriers are referred to as **subcarriers**. Each signal may use a different kind of modulation (e.g., DSB-SC, AM, SSB-SC, VSB-SC, or even frequency modulation or phase modulation). The modulated-signal

Figure 4.25 L-carrier hierarchical long-haul analog telephone frequency-division multiplexing system.



spectra may be separated by a small guard band to avoid interference and to facilitate signal separation at the receiver.

When all the modulated spectra have been added, we have a composite signal that may be considered to be a baseband signal to further modulate a radio-frequency (RF) carrier for the purpose of transmission.

At the receiver, the incoming signal is first demodulated by the RF carrier to retrieve the composite baseband, which is then bandpass-filtered to separate each modulated signal. Then each modulated signal is demodulated individually by an appropriate subcarrier to obtain all the basic baseband signals.

One simple example of FDM is the analog telephone long-haul system. There are two types of long-haul telephone carrier systems: the legacy analog L-carrier hierarchy systems and the digital T-carrier hierarchy systems in North America (or the E-carrier in Europe). Both were standardized by the predecessor of the International Telecommunications Union known (before 1992) as the CCITT (Comité Consultatif International Téléphonique et Télégraphique). We will first describe the analog telephone hierarchy that utilizes FDM and SSB modulation here and defer the digital hierarchy discussion until Chapter 6.

In the analog L-carrier hierarchy,[2] each voice channel is modulated using SSB+C. Twelve voice channels form a basic channel **group** occupying the bandwidth between 60 and 108 kHz. As shown in Fig. 4.25, each user channel uses LSB, and frequency division multiplexing (FDM) is achieved by maintaining the channel carrier separation of 4 kHz.

Further up the hierarchy,[3] five groups form a **supergroup** via FDM. Multiplexing 10 supergroups generates a **mastergroup** and multiplexing 6 supergroups forms a **jumbo group**, which consists of 3600 voice channels over a frequency band of 16.984 MHz in the L4 system. At each level of the hierarchy from the supergroup, additional frequency gaps are provided for interference reduction and for inserting pilot frequencies. The multiplexed signal can be fed into the baseband input of a microwave radio channel or directly into a coaxial transmission system.

4.8 PHASE-LOCKED LOOP AND APPLICATIONS

Phase-Locked Loop (PLL)

The **phase-locked loop** is a very important device typically used to track the phase and the frequency of the carrier component of an incoming signal. It is, therefore, a useful device for

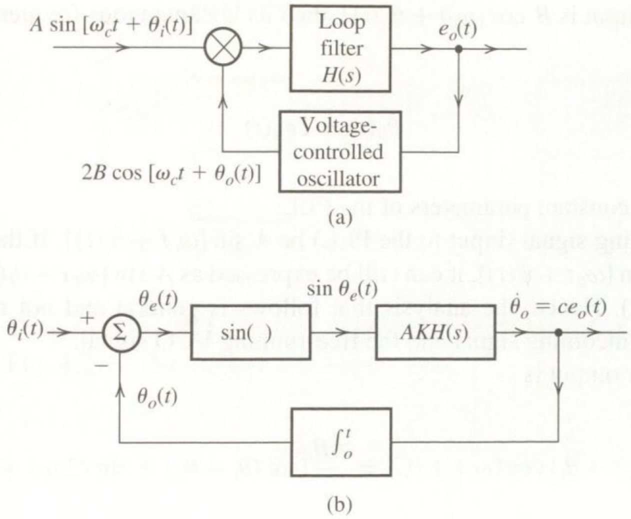

$A \sin[\omega_c t + \theta_i(t)]$
Loop filter $H(s)$
$e_o(t)$
Voltage-controlled oscillator
$2B \cos[\omega_c t + \theta_o(t)]$
(a)
$\sin \theta_e(t)$
$\theta_e(t)$
$\dot{\theta}_o = ce_o(t)$
$\theta_i(t)$
+
−
Σ
sin( )
$AKH(s)$
$\theta_o(t)$
$\int_o^t$
(b)

Figure 4.26 Phase-locked loop and its equivalent circuit.

the synchronous demodulation of AM signals with a suppressed carrier or with a little carrier (pilot). It can also be used for the demodulation of angle-modulated signals, especially under low SNR conditions. It also has important applications in a number of clock recovery systems including timing recovery in digital receivers. For these reasons, the PLL plays a key role in nearly every modern digital and analog communication system.

A PLL has three basic components:

1. A voltage-controlled oscillator (VCO).
2. A multiplier, serving as a phase detector (PD) or a phase comparator.
3. A loop filter $H(s)$.

Basic PLL Operation

The operation of the PLL is similar to that of a feedback system (Fig. 4.26a). In a typical feedback system, the feedback signal tends to follow the input signal. If the feedback signal is not equal to the input signal, the difference (known as the error) will change the feedback signal until it is close to the input signal. A PLL operates on a similar principle, except that the quantity fed back and compared is not the amplitude, but the phase. The VCO adjusts its own frequency such that its frequency and phase can track those of the input signal. At this point, the two signals are synchronous (except for a possible difference of a constant phase).

The **voltage-controlled oscillator (VCO)** is an oscillator whose frequency can be linearly controlled by an input voltage. If a VCO input voltage is $e_o(t)$, its output is a sinusoid with instantaneous (angular) frequency given by

$$\omega(t) = \omega_c + ce_o(t) \tag{4.30}$$

where c is a constant of the VCO and ω_c is the **free-running frequency** of the VCO [when $e_o(t) = 0$]. The multiplier output is further low-pass-filtered by the loop filter and then applied to the input of the VCO. This voltage changes the frequency of the oscillator and keeps the loop **locked** by forcing the VCO output to track the phase (and hence the frequency) of the input sinusoid.

If the VCO output is $B\cos[\omega_c t + \theta_o(t)]$, then its instantaneous frequency is $\omega_c + \dot{\theta}_o(t)$. Therefore,

$$\dot{\theta}_o(t) = c e_o(t) \tag{4.31}$$

where c and B are constant parameters of the PLL.

Let the incoming signal (input to the PLL) be $A\sin[\omega_c t + \theta_i(t)]$. If the incoming signal happens to be $A\sin[\omega_o t + \psi(t)]$, it can still be expressed as $A\sin[\omega_c t + \theta_i(t)]$, where $\theta_i(t) = (\omega_o - \omega_c)t + \psi(t)$. Hence, the analysis that follows is general and not restricted to equal frequencies of the incoming signal and the free-running VCO signal.

The multiplier output is

$$AB\sin(\omega_c t + \theta_i)\cos(\omega_c t + \theta_o) = \frac{AB}{2}[\sin(\theta_i - \theta_o) + \sin(2\omega_c t + \theta_i + \theta_o)]$$

The sum frequency term is suppressed by the loop filter. Hence, the effective input to the loop filter is $\frac{1}{2}AB\sin[\theta_i(t) - \theta_o(t)]$. If $h(t)$ is the unit impulse response of the loop filter,

$$\begin{aligned} e_o(t) &= h(t) * \frac{1}{2}AB\sin[\theta_i(t) - \theta_o(t)] \\ &= \frac{1}{2}AB\int_0^t h(t-x)\sin[\theta_i(x) - \theta_o(x)]\,dx \end{aligned} \tag{4.32}$$

Substituting Eq. (4.32) into Eq. (4.31) and letting $K = \frac{1}{2}cB$ lead to

$$\dot{\theta}_o(t) = AK\int_0^t h(t-x)\sin\theta_e(x)\,dx \tag{4.33}$$

where $\theta_e(t)$ is the phase error, defined as

$$\theta_e(t) = \theta_i(t) - \theta_o(t)$$

These equations [along with Eq. (4.31)] immediately suggest a model for the PLL, as shown in Fig. 4.26b.

The PLL design requires careful selection of the loop filter $H(s)$ and the loop gain AK. Different loop filters can enable the PLL to capture and track input signals with different types of frequency variation. On the other hand, the loop gain can affect the range of the trackable frequency variation.

Small-Error PLL Analysis

In small-error PLL analysis, $\sin\theta_e \simeq \theta_e$, and the block diagram in Fig. 4.26b reduces to the linear (time-invariant) system shown in Fig. 4.27a. Straightforward feedback analysis gives

$$\frac{\Theta_o(s)}{\Theta_i(s)} = \frac{AKH(s)/s}{1 + [AKH(s)/s]} = \frac{AKH(s)}{s + AKH(s)} \tag{4.34}$$

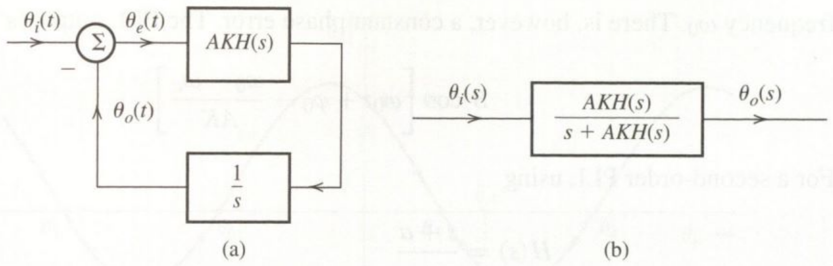


Figure 4.27 Equivalent circuits of a linearized PLL.

Therefore, the PLL acts as a filter with transfer function $AKH(s)/[s + AKH(s)]$, as shown in Fig. 4.27b. The error $\Theta_e(s)$ is given by

$$\Theta_e(s) = \Theta_i(s) - \Theta_o(s) = \left[1 - \frac{\Theta_o(s)}{\Theta_i(s)}\right]\Theta_i(s)$$

$$= \frac{s}{s + AKH(s)}\Theta_i(s) \tag{4.35}$$

One of the important applications of the PLL is in the acquisition of the frequency and the phase for the purpose of synchronization. Let the incoming signal be $A\ \sin(\omega_0 t + \varphi_0)$. We wish to generate a local signal of frequency ω_0 and phase* φ_0. Assuming the quiescent frequency of the VCO to be ω_c, the incoming signal can be expressed as $A\ \sin[\omega_c t + \theta_i(t)]$, where

$$\theta_i(t) = (\omega_0 - \omega_c)t + \varphi_0$$

and

$$\Theta_i(s) = \frac{\omega_0 - \omega_c}{s^2} + \frac{\varphi_0}{s}$$

Consider the special case of $H(s) = 1$. Substituting this equation into Eq. (4.35),

$$\Theta_e(s) = \frac{s}{s + AK}\left[\frac{\omega_0 - \omega_c}{s^2} + \frac{\varphi_0}{s}\right]$$

$$= \frac{(\omega_o - \omega_c)/AK}{s} - \frac{(\omega_0 - \omega_c)/AK}{s + AK} + \frac{\varphi_0}{s + AK}$$

Hence,

$$\theta_e(t) = \left[\frac{(\omega_0 - \omega_c)}{AK}\left(1 - e^{-AKt}\right) + \varphi_0 e^{-AKt}\right]u(t) \tag{4.36a}$$

Observe that

$$\lim_{t\to\infty} \theta_e(t) = \frac{\omega_0 - \omega_c}{AK} \tag{4.36b}$$

Hence, after the transient dies (in about $4/AK$ seconds), the phase error maintains a constant value of $(\omega_0 - \omega_c)/AK$. This means that the PLL frequency eventually equals the incoming frequency ω_0. There is, however, a constant phase error. The PLL output is

---

* With a difference $\pi/2$.

$$B \cos\left[\omega_0 t + \varphi_0 - \frac{\omega_0 - \omega_c}{AK}\right]$$

For a second-order PLL using

$$H(s) = \frac{s+a}{s} \tag{4.37a}$$

$$\Theta_e(s) = \frac{s}{s + AKH(s)}\Theta_i(s) \tag{4.37b}$$

$$= \frac{s^2}{s^2 + AK(s+a)}\left[\frac{\omega_0 - \omega_c}{s^2} + \frac{\varphi_0}{s}\right]$$

The final value theorem directly yields,[4]

$$\lim_{t\to\infty} \theta_e(t) = \lim_{s\to 0} s\Theta_e(s) = 0 \tag{4.38}$$

In this case, the PLL eventually acquires both the frequency and the phase of the incoming signal.

Using small-error analysis, it can be shown that a first-order loop cannot track an incoming signal whose instantaneous frequency varies linearly with time. Such a signal can be tracked within a constant phase (constant phase error) by using a second-order loop [Eq. (4.37)], and it can be tracked with zero phase error by using a third-order loop.[5]

It must be remembered that the preceding analysis assumes a linear model, which is valid only when $\theta_e(t) \ll \pi/2$. This means the frequencies ω_0 and ω_c must be very close for this analysis to be valid. For a general case, one must use the nonlinear model in Fig. 4.26b. For such an analysis, the reader is referred to Viterbi,[5] Gardner,[6] or Lindsey.[7]

First-Order Loop Analysis

Here we shall use the nonlinear model in Fig. 4.26b, but for the simple case of $H(s) = 1$. For this case $h(t) = \delta(t)$,* and Eq. (4.33) gives

$$\dot{\theta}_o(t) = AK \sin\theta_e(t)$$

Because $\theta_e = \theta_i - \theta_o$,

$$\dot{\theta}_e = \dot{\theta}_i - AK \sin\theta_e(t) \tag{4.39}$$

Let us here consider the problem of frequency and phase acquisition. Let the incoming signal be $A \sin(\omega_0 t + \varphi_0)$ and let the VCO have a quiescent frequency ω_c. Hence,

$$\theta_i(t) = (\omega_0 - \omega_c)t + \varphi_0$$

and

$$\dot{\theta}_e = (\omega_0 - \omega_c) - AK \sin\theta_e(t) \tag{4.40}$$

---

* Actually $h(t) = 2B \operatorname{sinc}(2\pi Bt)$, where B is the bandwidth of the loop filter. This is a low-pass, narrowband filter, which suppresses the high-frequency signal centered at $2\omega_c$. This makes $H(s) = 1$ over a low-pass narrow band of B Hz.

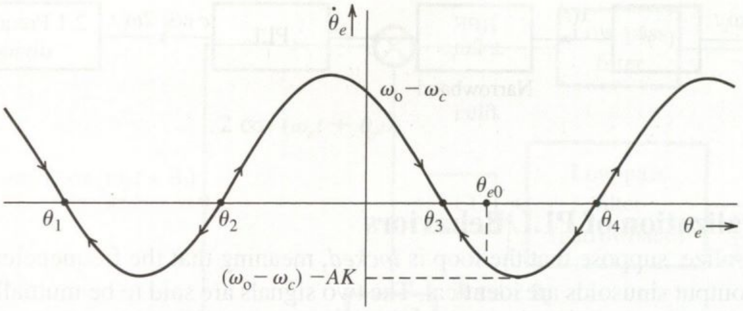


Figure 4.28 Trajectory of a first-order PLL.

For a better understanding of PLL behavior, we use Eq. (4.40) to sketch $\dot{\theta}_e$ vs. θ_e. Equation (4.40) shows that $\dot{\theta}_e$ is a vertically shifted sinusoid, as shown in Fig. 4.28. To satisfy Eq. (4.40), the loop operation must stay along the sinusoidal trajectory shown in Fig. 4.28. When $\dot{\theta}_e = 0$, the system is in equilibrium, because at these points, θ_e stops varying with time. Thus $\theta_e = \theta_1$, θ_2, θ_3, and θ_4 are all equilibrium points.

If the initial phase error $\theta_e(0) = \theta_{e0}$ (Fig. 4.28), then $\dot{\theta}_e$ corresponding to this value of θ_e is negative. Hence, the phase error will start decreasing along the sinusoidal trajectory until it reaches the value θ_3, where equilibrium is attained. Hence, in steady state, the phase error is a constant θ_3. This means the loop is in frequency lock; that is, the VCO frequency is now ω_0, but there is a phase error of θ_3. Note, however, that if $|\omega_0 - \omega_c| > AK$, there are no equilibrium points in Fig. 4.28, the loop never achieves lock, and θ_e continues to move along the trajectory forever. Hence, this simple loop can achieve phase lock provided the incoming frequency ω_0 does not differ from the quiescent VCO frequency ω_c by more than AK.

In Fig. 4.28, several equilibrium points exist. Half of these points, however, are unstable equilibrium points, meaning that a slight perturbation in the system state will move the operating point farther away from these equilibrium points. Points θ_1 and θ_3 are stable points because any small perturbation in the system state will tend to bring it back to these points. Consider, for example, the point θ_3. If the state is perturbed along the trajectory toward the right, $\dot{\theta}_e$ is negative, which tends to reduce θ_e and bring it back to θ_3. If the operating point is perturbed from θ_3 toward the left, $\dot{\theta}_e$ is positive, θ_e will tend to increase, and the operating point will return to θ_3. On the other hand, at point θ_2 if the point is perturbed toward the right, $\dot{\theta}_e$ is positive, and θ_e will increase until it reaches θ_3. Similarly, if at θ_2 the operating point is perturbed toward the left, $\dot{\theta}_e$ is negative, and θ_e will decrease until it reaches θ_1. Hence, θ_2 is an unstable equilibrium point. The slightest disturbance, such as noise, will dislocate it either to θ_1 or to θ_3. In a similar way, we can show that θ_4 is an unstable point and that θ_1 is a stable equilibrium point.

The equilibrium point θ_3 occurs where $\dot{\theta}_e = 0$. Hence, from Eq. (4.40),

$$\theta_3 = \sin^{-1}\frac{\omega_0 - \omega_c}{AK}$$

If $\theta_3 \ll \pi/2$, then

$$\theta_3 \simeq \frac{\omega_0 - \omega_c}{AK}$$

which agrees with our previous result of the small-error analysis [Eq. (4.36b)].

The first-order loop suffers from the fact that it has a constant phase error. Moreover, it can acquire frequency lock only if the incoming frequency and the VCO quiescent frequency differ by not more than AK rad/s. Higher order loops overcome these disadvantages, but they create a new problem of stability. More detailed analysis can be found in the book by Gardner.[6]

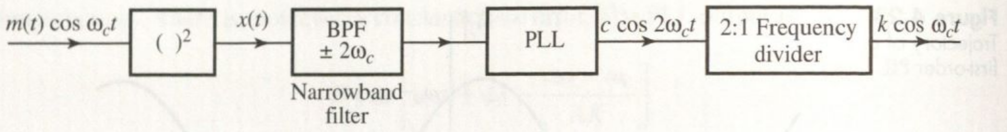


Figure 4.29 Generation of coherent demodulation carrier using signal squaring.

Generalization of PLL Behaviors

To generalize, suppose that the loop is *locked*, meaning that the frequencies of both the input and the output sinusoids are identical. The two signals are said to be mutually **phase coherent** or **in phase lock**. The VCO thus tracks the frequency and the phase of the incoming signal. A PLL can track the incoming frequency only over a finite range of frequency shift. This range is called the **hold-in** or **lock** range. Moreover, if initially the input and output frequencies are not close enough, the loop may not acquire lock. The frequency range over which the input will cause in the loop to lock is called the **pull-in** or **capture** range. Also if the input frequency changes too rapidly, the loop may not lock.

If the input sinusoid is noisy, the PLL not only tracks the sinusoid, but also cleans it up. The PLL can also be used as a frequency modulation (FM) demodulator and frequency synthesizer, as shown later in the next chapter. Frequency multipliers and dividers can also be built using PLL. The PLL, being a relatively inexpensive integrated circuit, has become one of the most frequently used communication circuits.

In space vehicles, because of the Doppler shift and oscillator drift, the frequency of the received signal has much uncertainty. The Doppler shift of the carrier itself could be as high as ±75 kHz, whereas the desired modulated signal band may be just 10 Hz. To receive such a signal by conventional receivers would require a filter of bandwidth 150 kHz, when the desired signal has a bandwidth of only 10 Hz. This would cause an undesirable increase (by a factor of 15,000) in the noise received because the noise power is proportional to the bandwidth. The PLL proves convenient here because it tracks the received frequency continuously, and the filter bandwidth required is only 10 Hz.

Carrier Acquisition in DSB-SC

We shall now discuss two methods of carrier regeneration using PLL at the receiver in DSB-SC: signal squaring and the Costas loop.

Signal-Squaring Method: An outline of this scheme is given in Fig. 4.29. The incoming signal is squared and then passed through a narrow (high-Q) bandpass filter tuned to $2\omega_c$. The output of this filter is the sinusoid $k\cos 2\omega_c t$, with some residual unwanted signal. This signal is applied to a PLL to obtain a cleaner sinusoid of twice the carrier frequency, which is passed through a 2:1 frequency divider to obtain a local carrier in phase and frequency synchronism with the incoming carrier. The analysis is straightforward. The squarer output $x(t)$ is

$$x(t) = [m(t)\cos\omega_c t]^2 = \frac{1}{2}m^2(t) + \frac{1}{2}m^2(t)\cos 2\omega_c t$$

Now $m^2(t)$ is a nonnegative signal, and therefore has a nonzero average value [in contrast to $m(t)$, which generally has a zero average value]. Let the average value, which is the dc component of $m^2(t)/2$, be k. We can now express $m^2(t)/2$ as

$$\frac{1}{2}m^2(t) = k + \phi(t)$$

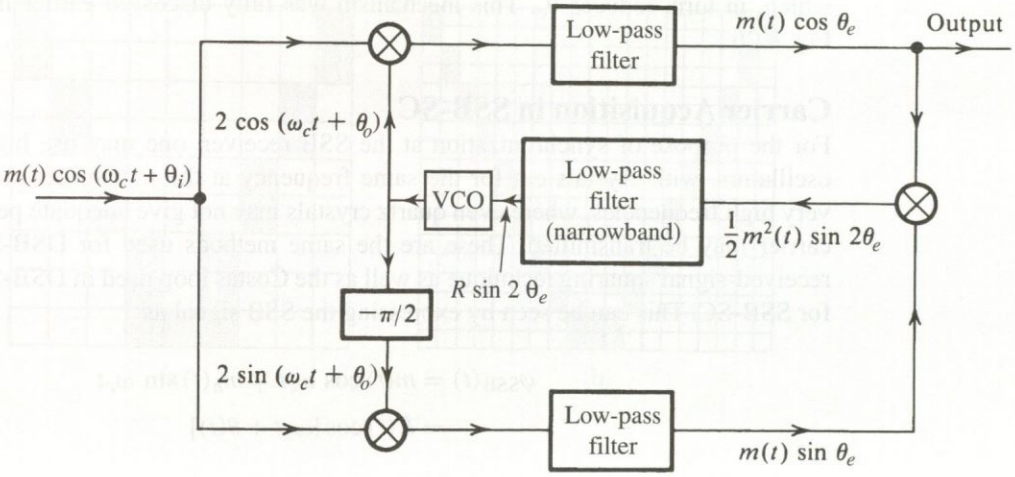

Low-pass filter
$m(t) \cos \theta_e$
Output
$2 \cos (\omega_c t + \theta_o)$
$m(t) \cos (\omega_c t + \theta_i)$
VCO
Low-pass filter (narrowband)
$\frac{1}{2} m^2(t) \sin 2\theta_e$
$R \sin 2\theta_e$
$-\pi/2$
$2 \sin (\omega_c t + \theta_o)$
Low-pass filter
$m(t) \sin \theta_e$

Figure 4.30 Costas phase-locked loop for the generation of a coherent demodulation carrier.

where $\phi(t)$ is a zero mean baseband signal [$m^2(t)/2$ minus its dc component]. Thus,

$$
\begin{aligned}
x(t) &= \frac{1}{2}m^2(t) + \frac{1}{2}m^2(t) \cos 2\omega_c t \\
&= \frac{1}{2}m^2(t) + k \cos 2\omega_c t + \phi(t) \cos 2\omega_c t.
\end{aligned}
$$

The bandpass filter is a narrowband (high-Q) filter tuned to frequency $2\omega_c$. It completely suppresses the signal $m^2(t)$, whose spectrum is centered at $\omega = 0$. It also suppresses most of the signal $\phi(t) \cos 2\omega_c t$. This is because although this signal spectrum is centered at $2\omega_c$, it has zero (infinitesimal) power at $2\omega_c$ since $\phi(t)$ has a zero dc value. Moreover this component is distributed over the band of $4B$ Hz centered at $2\omega_c$. Hence, very little of this signal passes through the narrowband filter.* In contrast, the spectrum of $k \cos 2\omega_c t$ consists of impulses located at $\pm 2\omega_c$. Hence, all its power is concentrated at $2\omega_c$ and will pass through. Thus, the filter output is $k \cos 2\omega_c t$ plus a small undesired residue from $\phi(t) \cos 2\omega_c t$. This residue can be suppressed by using a PLL, which tracks $k \cos 2\omega_c t$. The PLL output, after passing through a 2:1 frequency divider, yields the desired carrier. One qualification is in order. Because the incoming signal sign is lost in the squarer, we have sign ambiguity (or phase ambiguity of π) in the carrier generated. This is immaterial for analog signals. For a digital baseband signal, however, the carrier sign is essential, and this method, therefore, must be modified.

Costas Loop: Yet another scheme for generating a local carrier, proposed by Costas,[8] is shown in Fig. 4.30. The incoming signal is $m(t) \cos (\omega_c t + \theta_i)$. At the receiver, a VCO generates the carrier $\cos (\omega_c t + \theta_o)$. The phase error is $\theta_e = \theta_i - \theta_o$. Various signals are indicated in Fig. 4.30. The two low-pass filters suppress high-frequency terms to yield $m(t) \cos \theta_e$ and $m(t) \sin \theta_e$, respectively. These outputs are further multiplied to give $m^2(t) \sin 2\theta_e$. When this is passed through a narrowband, low-pass filter, the output is $R \sin 2\theta_e$, where R is the dc component of $m^2(t)/2$. The signal $R \sin 2\theta_e$ is applied to the input of a VCO with quiescent frequency ω_c. The input $R \sin 2\theta_e$ increases the output frequency, which, in turn, reduces θ_e. This mechanism was fully discussed earlier in connection with Fig. 4.26.

---

* This will also explain why we cannot extract the carrier directly from $m(t) \cos \omega_c t$ by passing it through a narrowband filter centered at ω_c. The reason is that the power of $m(t) \cos \omega_c t$ at ω_c is zero because $m(t)$ has no dc component [the average value of $m(t)$ is zero].

Carrier Acquisition in SSB-SC

For the purpose of synchronization at the SSB receiver, one may use highly stable crystal oscillators, with crystals cut for the same frequency at the transmitter and the receiver. At very high frequencies, where even quartz crystals may not give adequate performance, a pilot carrier may be transmitted. These are the same methods used for DSB-SC. However, the received-signal squaring technique as well as the Costas loop used in DSB-SC cannot be used for SSB-SC. This can be seen by expressing the SSB signal as

$$\begin{aligned}\varphi_{\text{SSB}}(t) &= m(t)\cos\omega_c t \mp m_h(t)\sin\omega_c t\\ &= E(t)\cos[\omega_c t + \theta(t)]\end{aligned}$$

where

$$\begin{aligned}E(t) &= \sqrt{m^2(t) + m_h^2(t)}\\ \theta(t) &= \tan^{-1}\left[\frac{\pm m_h(t)}{m(t)}\right]\end{aligned}$$

Squaring this signal yields

$$\begin{aligned}\varphi_{\text{SSB}}^2(t) &= E^2(t)\cos^2[\omega_c t + \theta(t)]\\ &= \frac{E^2(t)}{2}\{1 + \cos[2\omega_c t + 2\theta(t)]\}\end{aligned}$$

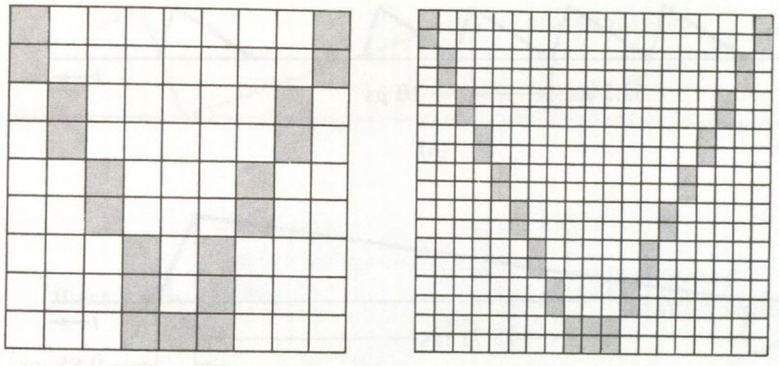
The signal $E^2(t)$ is eliminated by a bandpass filter. Unfortunately, the remaining signal is not a pure sinusoid of frequency $2\omega_c$ (as was the case for DSB). There is nothing we can do to remove the time-varying phase $2\theta(t)$ from this sinusoid. Hence, for SSB, the squaring technique does not work. The same argument can be used to show that the Costas loop will not work either. These conclusions also apply to VSB signals.

4.9 NTSC TELEVISION BROADCASTING SYSTEM

Television is by far the most successful of the commercial mass media. Analog TV broadcasting has a long history. As early as 1928, C. F. Jenkins launched his station W3XK for television broadcasting. The NTSC (National Television System Committee) first issued its standard for black-and-white TV broadcasting in 1941 and its NTSC color standard in 1953. Although high-definition TV (HDTV) is poised to make the analog NTSC system obsolete, much can be learned about the AM system, and the FM system in the next chapter, from a good understanding of how the NTSC system functions.

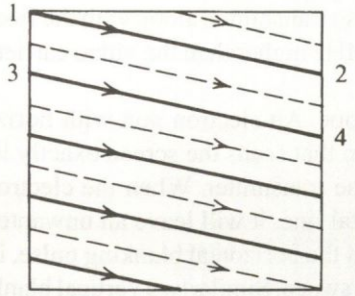
In television, the central problem is the transmission of visual images by electrical signals. The image, or picture, can be thought of as a frame subdivided into several small squares, known as picture elements (or pixels). A large number of picture elements in a given image means clearer reproduction (better resolution) at the receiver (see Fig. 4.31). The information of the entire picture is transmitted by transmitting an electric signal proportional to the brightness level of the pixels taken in a certain sequence. We start from the upper left-hand corner with

Figure 4.31 Effect of the number of picture elements on resolution. (a) Poor resolution (b) Better resolution.



(a) (b)

Figure 4.32 Scanning pattern (raster).

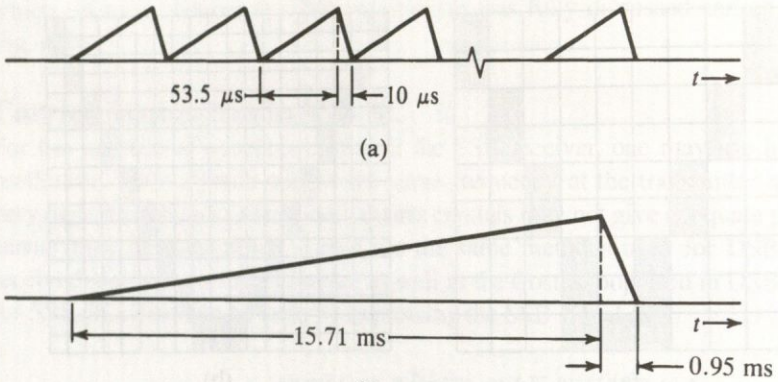


element number 1 and scan the first row of pixels (Fig. 4.32). Then we come back to the start of the second row, scan the second row, and continue this way until we have finished the last row. The electric signal thus generated during the entire scanning interval has the information contained in the picture.

The image is furnished by the television camera tube. There exist a variety of camera tubes. The image orthicon is one example. In this tube, the optical system generates a focused image on a photo cathode, which eventually produces an electrically **charged image** on another surface, known as the **target mosaic**. What this means is that every point on the target mosaic surface acquires a positive electric charge proportional to the brightness of the image. Thus, instead of a light image, we have a charge image. An electron gun now scans the target mosaic surface with an electron beam in the manner shown in Fig. 4.32. The beam is controlled by a set of voltages across horizontal and vertical deflection plates. Periodic sawtooth signals (Fig. 4.33) are applied to these plates. The beam scans the horizontal line 1–2 in 53.5 μs and quickly flies back, in 10 μs, to the left to point 3 and scans line 3–4, and so on. On the target mosaic, where there is a high positive charge (corresponding to a higher brightness level), more electrons from the beam will be absorbed, and the return beam will have fewer electrons, giving a smaller current. Areas corresponding to darker elements (less positive charge) will return a large current. The scanning lines are not perfectly horizontal but have a small downward slope, since during the horizontal deflection the beam is also continuously deflected downward owing to a slower vertical deflection signal (Fig. 4.33b). When all the horizontal lines have been scanned, the vertical deflection signal goes to zero, which means the beam goes back to point 1 again and is ready to start the next frame.

Scanning is continuous at a rate of 60 picture frames per second. The electrical signal thus generated is a video signal corresponding to the visual image. With some modifications (to be

Figure 4.33
(a) Horizontal deflection signal.
(b) Vertical deflection signal.



discussed later), the video signal modulates the video carrier of frequency f_c in the form of VSB-AM (Fig. 4.24). This video carrier is also transmitted along with the frequency-modulated audio carrier of frequency f_α, which is 4.5 MHz higher than the video carrier frequency f_c, that is, $f_\alpha = f_c + 4.5$ MHz.

The receiver is similar to an oscilloscope. An electron gun with horizontal and vertical deflection plates generates an electron beam that scans the screen exactly in the same pattern and in synchronism with the scanning at the transmitter. When the electron beam flies back horizontally after completing each horizontal line, it will leave an unwanted flyback trace on the screen unless a blanking pulse, known as the horizontal blanking pulse, is added during the flyback interval at the end of each horizontal sweep. Similarly, a vertical blanking pulse is added at the end of each vertical sweep to eliminate the unwanted vertical retrace. These blanking pulses are added at the transmitter itself. We also need to add scan-synchronization information at the transmitter. This is done by adding a large pulse to each blanking pulse. A typical video signal is shown in Fig. 4.34. It is evident that over the entire flyback interval, the blanking pulse (at the black level) will eliminate the trace. Similarly, vertical blanking and synchronizing pulses, which are much wider than the corresponding horizontal pulses, are added to the video signal at the end of each vertical sweep. The video signal now VSB-modulates the carrier. We also add a carrier at this point (see Fig. 4.24). This VSB+C signal is transmitted along with the frequency-modulated audio signal. The transmitter block diagram is shown in Fig. 4.35a, the receiver block diagram in Fig. 4.35b. This is a superheterodyne receiver (see Example 4.2). The reasons for using a superheterodyne receiver will be discussed later in Sec. 5.6. The converter (a mixer) shifts the entire spectrum (video as well as frequency-modulated audio) to the IF frequency. This signal is now amplified and envelope-detected. The audio signal is still of the frequency-modulated form with a carrier of 4.5 MHz. It is separated and demodulated. The video signal is amplified. Synchronizing pulses are separated and applied to the vertical and horizontal sweep generators. The video signal is clamped to the blanking pulses (dc restoration) and then applied to the picture tube.

Bandwidth Considerations

The number of horizontal lines used in the NTSC standard is 495 per frame. The time required for vertical retrace at the end of the scan is equivalent to that required for 30 horizontal lines. Hence, each frame is considered to have a total of 525 lines,* out of which only 495 are active. Images must be transmitted in a rapid succession of frames to create the illusion of continuity

---

* In Europe, a total of 625 horizontal lines is used.

Figure 4.34 Television video signal.

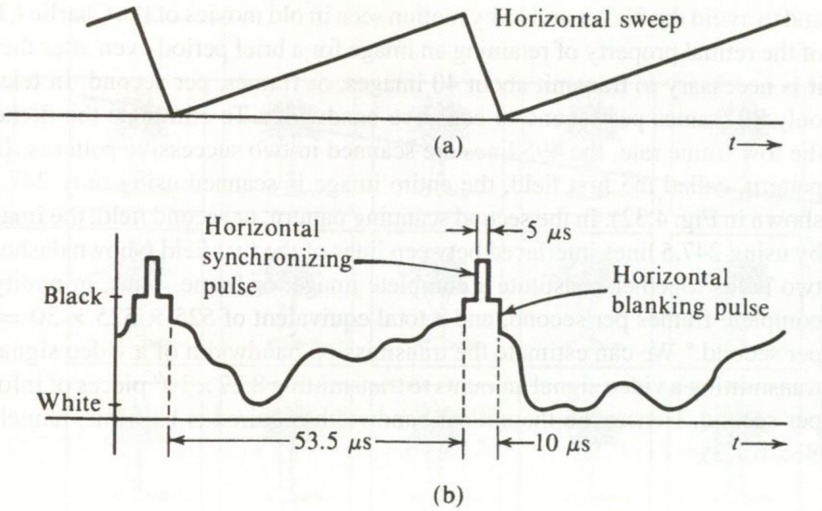
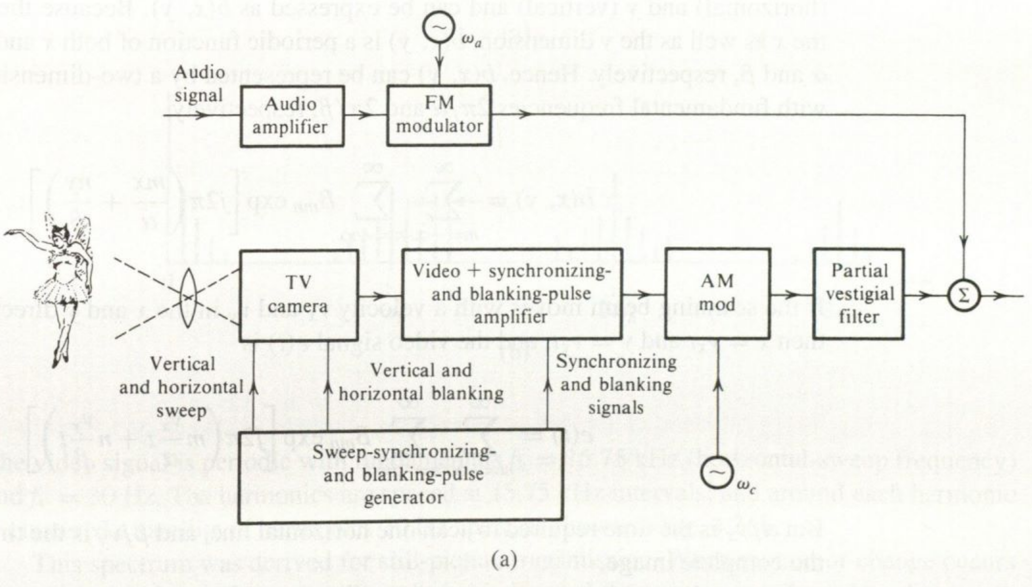


Figure 4.35 (a) Television transmitter. (b) Television receiver.


ω_a
Audio
signal
Audio
amplifier
FM
modulator
TV
camera
Video + synchronizing-
and blanking-pulse
amplifier
AM
mod
Partial
vestigial
filter
Σ
Vertical
and horizontal
sweep
Vertical and
horizontal blanking
Synchronizing
and blanking
signals
ω_c
Sweep-synchronizing-
and blanking-pulse
generator
(a)
Synchronization
separator
Sweep
generator
Received
TV carrier
RF
shaping
filter &
amplifier
Frequency
converter
IF
amplifier
Envelope
detector
Video amplifiers
& synchronization
separator
Dc
restorer
Picture tube
$\omega_{LO} = \omega_c + \omega_{IF}$
FM
demodulator
Audio
amplifier
Speaker
(b)

and to avoid the flicker and jerky motion seen in old movies of the Charlie Chaplin era. Because of the retinal property of retaining an image for a brief period even after the object is removed, it is necessary to transmit about 40 images, or frames, per second. In television we transmit only 30 frames per second to conserve bandwidth. To eliminate the flicker effect caused by the low frame rate, the 495 lines are scanned in two successive patterns. In the first scanning pattern, called the first field, the entire image is scanned using only 247.5 lines (solid lines shown in Fig. 4.32). In the second scanning pattern, or second field, the image is scanned again by using 247.5 lines interlaced between lines of the first field (shown dashed in Fig. 4.32). The two fields together constitute a complete image, or frame. Thus, in reality there are only 30 complete frames per second, and a total equivalent of $525 \times 525 \times 30 = 8.27 \times 10^6$ pixels per second.* We can estimate the transmission bandwidth of a video signal by observing that transmitting a video signal amounts to transmitting 8.27×10^6 pieces of information (or pulses) per second. Hence, the theoretical bandwidth required is half this, namely, 4.135 MHz (see Sec. 6.1.3).

Video Spectrum

To begin, consider a simple case of transmission of a still image. The scanning procedure discussed earlier is equivalent to scanning an array of the same image repeating itself in both dimensions, as shown in Fig. 4.36a. The brightness level b for this figure is a function of x (horizontal) and y (vertical) and can be expressed as $b(x, y)$. Because the picture repeats in the x as well as the y dimension, $b(x, y)$ is a periodic function of both x and y, with periods of α and β, respectively. Hence, $b(x, y)$ can be represented by a two-dimensional Fourier series with fundamental frequencies $2\pi/\alpha$ and $2\pi/\beta$, respectively,

$$b(x, y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} B_{mn} \exp\left[j2\pi\left(\frac{mx}{\alpha} + \frac{ny}{\beta}\right)\right] \tag{4.41a}$$

If the scanning beam moves with a velocity v_x and v_y in the x and y directions, respectively, then $x = v_x t$ and $y = v_y t$, and the video signal $e(t)$ is

$$e(t) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} B_{mn} \exp\left[j2\pi\left(m\frac{v_x}{\alpha}t + n\frac{v_y}{\beta}t\right)\right] \tag{4.41b}$$

But α/v_x is the time required to scan one horizontal line, and β/v_y is the time required to scan the complete image,

$$\frac{\alpha}{v_x} = \frac{1}{30(525)} \quad \text{and} \quad \frac{\beta}{v_y} = \frac{1}{30}$$

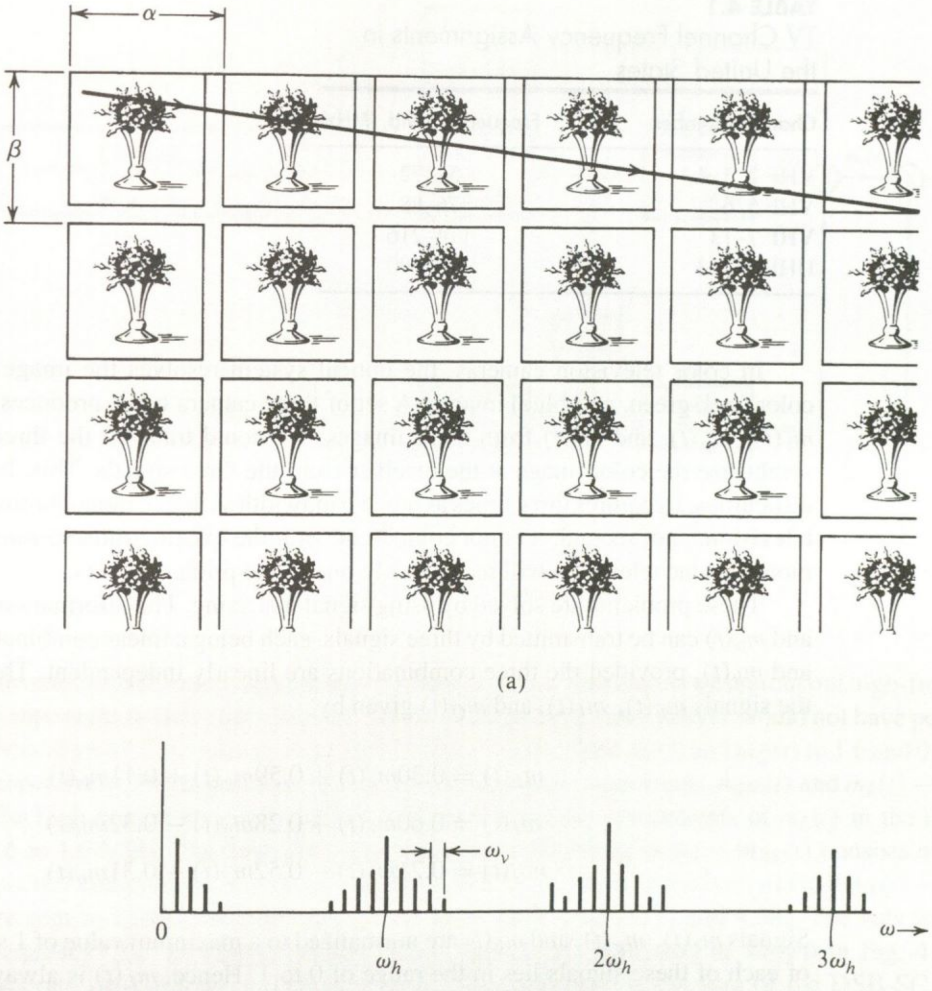
and

$$e(t) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} B_{mn} \exp\left[j2\pi(15{,}750m + 30n)t\right] \tag{4.41c}$$

---

[*] Actually, the ratio of the image width to the image height (aspect ratio) is 4:3. Hence the number of picture elements will increase by a factor of 4/3. But this factor is almost canceled out because the scanning pattern does not align perfectly with the checkerboard pattern in Fig. 4.31, thus reducing the resolution by a factor of 0.70 (the Kell factor).

Figure 4.36
(a) Model for scanning process using doubly periodic image fields.
(b) Spectrum of the monochrome video signal.



The video signal is periodic with fundamentals $f_h = 15.75$ kHz (horizontal-sweep frequency) and $f_v = 30$ Hz. The harmonics are spaced at 15.75 kHz intervals, and around each harmonic is clustered a satellite of harmonics 30 Hz apart, as shown in Fig. 4.36b.

This spectrum was derived for still-picture transmission. When motion or change occurs from frame to frame, $b(x, y)$ will not be periodic, and the spectrum will not be a line spectrum, but will have spreading or smearing. But empty spaces still exist between harmonics of f_h (15.75 kHz). We take advantage of these gaps to transmit the additional information of a color television signal over the same bandwidth.

The FCC allows a 6 MHz bandwidth for television broadcasting, with the frequency allocations as shown in Table 4.1.

Compatible Color Television (CCTV)

All colors can be synthesized by mixing the three primary colors—blue, yellow, and red—in the right amounts. In television, blue, green (the combination of blue and yellow), and red are used instead for the practical reason of the availability of phosphors that glow with these colors when excited by an electron beam.

TABLE 4.1
TV Channel Frequency Assignments in the United States

Channel Number	Frequency Band, (MHz)
VHF 2, 3, 4	54–72
VHF 5, 6	76–88
VHF 7–13	174–216
UHF 14–83	470–890

In color television cameras, the optical system resolves the image into three primary color (red, green, and blue) images. A set of three camera tubes produces three video signals $m_r(t)$, $m_g(t)$, and $m_b(t)$ from these images. We could transmit the three video signals and synthesize the color image at the receiver from the three signals. This, however, causes two difficulties. It requires three times as much bandwidth as that of monochrome (black-and-white) television, and, second, it is not compatible with the existing monochrome system because a monochrome television will receive only one of the primary colors.

These problems are solved by using signal matrixing. The information about $m_r(t), m_g(t)$, and $m_b(t)$ can be transmitted by three signals, each being a linear combination of $m_r(t), m_g(t)$, and $m_b(t)$, provided the three combinations are linearly independent. Thus, we can transmit the signals $m_L(t)$, $m_I(t)$, and $m_Q(t)$ given by

$$
\begin{aligned}
m_L(t) &= 0.30m_r(t) + 0.59m_g(t) + 0.11m_b(t) \\
m_I(t) &= 0.60m_r(t) + 0.28m_g(t) - 0.32m_b(t) \\
m_Q(t) &= 0.21m_r(t) - 0.52m_g(t) + 0.31m_b(t)
\end{aligned}
$$

Signals $m_r(t)$, $m_g(t)$, and $m_b(t)$ are normalized to a maximum value of 1 so that the amplitude of each of these signals lies in the range of 0 to 1. Hence, $m_L(t)$ is always positive, whereas $m_I(t)$ and $m_Q(t)$ are bipolar. The signal $m_L(t)$ is known as the **luminance** signal because it has been found that this particular combination of the three primary color signals closely matches the luminance of the conventional monochrome video signal. Hence, a black-and-white set need use only this signal for its operation.

The signals $m_I(t)$ and $m_Q(t)$ are known as the **chrominance** signals.* We could have chosen some other combinations instead of $m_I(t)$ and $m_Q(t)$. But these particular combinations are chosen because they use certain features of human color vision efficiently,[9] as explained next.

Multiplexing Luminance and Chrominance Signals: The luminance signal $m_L(t)$ is transmitted as a monochrome video signal occupying a bandwidth of 4.2 MHz. The chrominance signals $m_I(t)$ and $m_Q(t)$ also have the same bandwidth (namely, 4.2 MHz each). Subjective tests have shown, however, that the human eye does not perceive changes in

---

* These signals have an interesting interpretation in terms of the **hue** and the **saturation** of colors. Hue refers to the color, such as red, yellow, green, blue, or any color in between. Saturation, or color intensity, refers to the purity of the color. For example, a deep red has 100% saturation, but pink—which is a dilution of red with white—will have a lesser amount of saturation. Saturation is given by $\sqrt{m_I^2(t) + m_Q^2(t)}$, and hue is given by an angle $\tan^{-1}[m_Q(t)/m_I(t)]$. Each color has a certain hue, or angle. For example, red, blue, and green are at angles of 19°, 136°, and 242°, respectively.

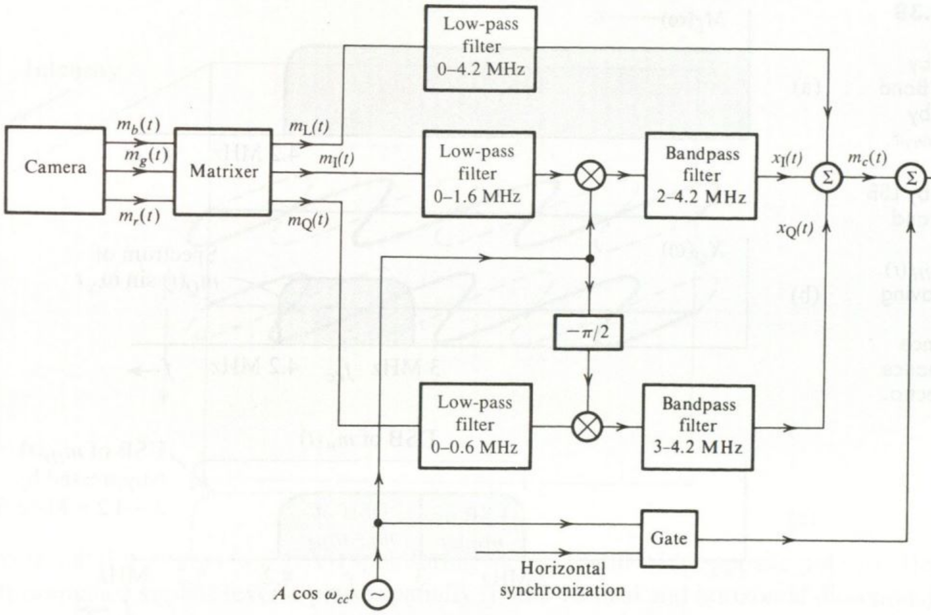

Low-pass filter 0–4.2 MHz
Camera
$m_b(t)$
$m_g(t)$
$m_r(t)$
Matrixer
$m_L(t)$
$m_I(t)$
$m_Q(t)$
Low-pass filter 0–1.6 MHz
Bandpass filter 2–4.2 MHz
$x_I(t)$
$m_c(t)$
$m_v(t)$
Σ
Σ
$x_Q(t)$
$-\pi/2$
Low-pass filter 0–0.6 MHz
Bandpass filter 3–4.2 MHz
Gate
Horizontal synchronization
$A \cos \omega_{cc} t$

Figure 4.37 Multiplexing luminance and chrominance signals.

chrominance (hue and saturation) over smaller areas. This means we can cut out high-frequency components without affecting the quality of the picture, since the eye would not have perceived them anyway. This enables us to limit the bandwidths of the $m_I(t)$ and $m_Q(t)$ to 1.6 and 0.6 MHz, respectively. The signal $m_I(t)$ is further split into two components, $m_{IH}(t)$ and $m_I(t) - m_{IH}(t)$. The high-frequency component $m_{IH}(t)$ consists of the components of $m_I(t)$ in the range of 0.6 to 1.6 MHz. The remaining low-frequency component $m_I(t) - m_{IH}(t)$ consists of all the spectral components of $m_I(t)$ in the range of 0 to 0.6 MHz. Signals $m_Q(t)$ and $m_I(t) - m_{IH}(t)$ are sent by QAM, whereas $m_{IH}(t)$ is sent by LSB (Figs. 4.37 and 4.38). The subcarrier has frequency* $f_{cc} = 3.583125$ MHz. These spectra are generated as shown in Fig. 4.37. We generate the DSB-SC spectrum of $m_I(t)$. This spectrum is the sum of the DSB-SC spectra of both its components $m_{IH}(t)$ and $m_I(t) - m_{IH}(t)$. This spectrum occupies a band of 2 to 5.2 MHz. However, the bandpass filter (2–4.2 MHz) suppresses the USB portion of $m_{IH}(t)$, leaving only the LSB spectrum of $m_{IH}(t)$ (Fig. 4.38). We still have the DSB-SC spectrum for $m_I(t) - m_{IH}(t)$. Thus, $x_I(t)$ consists of an LSB for $m_{IH}(t)$ and a DSB-SC for $m_I(t) - m_{IH}(t)$, and can be expressed as

$$\begin{aligned} x_I(t) &= \underbrace{[m_I(t) - m_{IH}(t)] \cos \omega_{cc} t}_{\text{DSB(QAM) for } m_I(t) - m_{IH}(t)} + \underbrace{m_{IH}(t) \cos \omega_{cc} t + m_{IH_h}(t) \sin \omega_{cc} t}_{\text{LSB for } m_{IH}(t)} \\ &= m_I(t) \cos \omega_{cc} t + m_{IH_h}(t) \sin \omega_{cc} t \end{aligned} \tag{4.42}$$

Moreover, the signal $x_Q(t) = m_Q(t) \sin \omega_{cc}(t)$. Hence, the composite multiplexed signal $m_c(t)$ is

$$m_c(t) = m_L(t) + m_Q(t) \sin \omega_{cc} t + m_I(t) \cos \omega_{cc} t + m_{IH_h}(t) \sin \omega_{cc} t$$

---

* $f_{cc} = 227.5 f_h = 227.5 \times 15.75$ kHz $= 3.583125$ MHz. Thus, f_{cc} lies midway between $227 f_h$ and $228 f_h$. This causes the chrominance signals' spectra to be shifted to gaps midway between harmonics of f_h (Fig. 4.38d). In practice, f_{cc} is made slightly smaller than $227.5 f_h$ ($f_{cc} = 3.579545$ MHz) to avoid an objectionable beat frequency with the audio carrier,[9] which lies 4.5 MHz above the picture carrier. Because $f_h = f_{cc}/227.5$, $f_h = 15.7326$ kHz, and the field repetition frequency is actually 59.94 rather than 60.

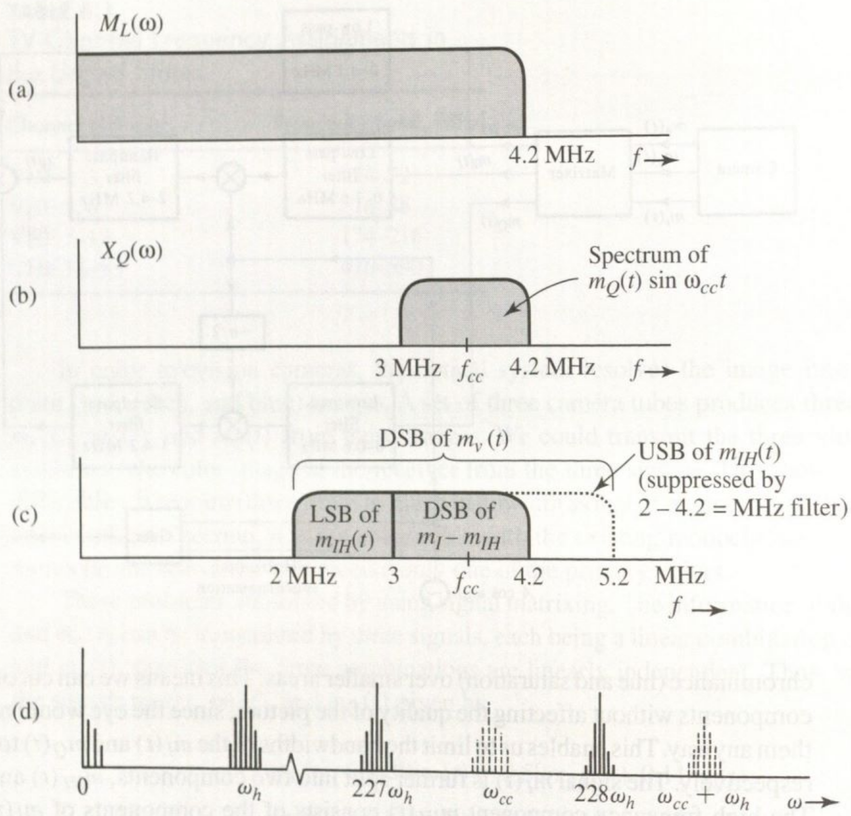


Figure 4.38 (a) Band occupied by $m_L(t)$. (b) Band occupied by $m_Q(t) \sin \omega_{cc} t$. (c) Bands occupied by LSB of $m_{IH}(t)$ and DSB of $m_I(t) - m_{IH}(t)$. (d) Interleaving of the chrominance and luminance signal spectra.

In addition, a sample of the subcarrier (color burst) is added to this multiplexed signal for frequency and phase synchronization of the locally generated subcarrier at the receiver. The color burst is added on the trailing edge of the horizontal blanking pulse. This composite video signal is now sent by VSB+C, as discussed in Sec. 4.5.

The Receiver: Because the CCTV system is required to be compatible with monochrome receivers, let us see what happens if we apply the signal $m_v(t)$ to a monochrome receiver. It may seem necessary to remove the chrominance signals from $m_v(t)$ before applying the signal to the picture tube. Fortunately, this step is not necessary, because the interference of the chrominance signals with the luminance signal, although present on the screen, is practically invisible to the human eye. This happens because of the way chrominance signals are interlaced in the frequency domain and because of the persistence of human vision that tends to average out brightness over time as well as space.

The chrominance signal is superimposed on the luminance signal. Figure 4.39 shows the nature of the chrominance signals. Recall that $\omega_{cc} = 227.5\omega_h$. During the traversal of one horizontal line, there will be 227.5 chrominance signal cycles. Hence, the chrominance signal changes continuously from positive to negative, and vice versa, in the horizontal direction. In addition, because there are 227.5 cycles in one line, if a chrominance signal begins with a positive cycle at the start of a line, it will end with a positive cycle at the end of the line. The next horizontal line will begin with a negative cycle of the chrominance signal (Fig. 4.39). Hence, in any given frame, the chrominance signal not only reverses its phase along the horizontal (x) direction but also reverses its phase along the vertical (y) direction (on the next horizontal line). But this is not all. During one field, the chrominance signal completes 227.5×525

Figure 4.39 Temporal and spatial phase reversals of chrominance signals: solid lines, first field; dashed lines, second field.

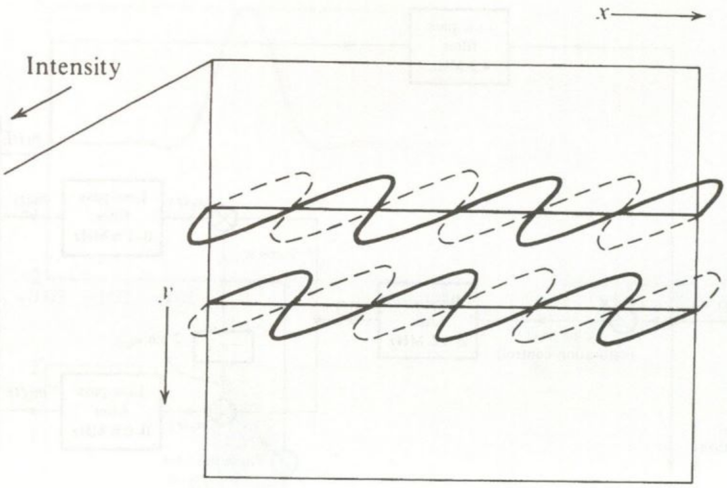


Figure 4.39 Temporal and spatial phase reversals of chrominance signals: solid lines, first field; dashed lines, second field.

cycles, and it returns to a given spot during the next field with opposite polarity. Hence, the chrominance signals reverse phase spatially (in the vertical and horizontal directions) as well as temporally at any given spot. Because the human eye is not sensitive to rapid variations in time or space, it can notice only space and time averages. This makes the chrominance signals practically invisible to the human eye. Thus the color signal is compatible with an unmodified monochrome receiver.

Demultiplexing: In a color receiver the received signal is demodulated exactly as in the monochrome case. This yields $m_v(t)$. This signal must now be demultiplexed to separate $m_L(t)$, $m_I(t)$, and $m_Q(t)$. The demultiplexing is shown in Fig. 4.40. The output of the 4.2 MHz filter contains $m_L(t)$, as well as modulated $m_I(t)$ and $m_Q(t)$ (Fig. 4.38). Because of the frequency interlacing discussed earlier, however, these signals are practically invisible. Hence, the output of the 4.2 MHz filter serves the function of $m_L(t)$. Next we use carriers in phase quadrature to demodulate $m_v(t)$. To determine the various signals in Fig. 4.40, we observe that the signal $z(t)$ in Fig. 4.40 consists of modulated $m_I(t)$ and $m_Q(t)$, plus the part of $m_L(t)$ in the band of 2 to 4.2 MHz. Let us denote this high-frequency component of $m_L(t)$ by $m_{LH}(t)$. Then,

$$z(t) = m_{LH}(t) + m_Q(t)\sin\omega_{cc}t + m_I(t)\cos\omega_{cc}t + m_{IH_h}(t)\sin\omega_{cc}t$$

Hence,

$$x_1(t) = 2m_{LH}(t)\cos\omega_{cc}t + m_Q(t)\sin 2\omega_{cc}t + m_I(t)(1+\cos 2\omega_c t) + m_{IH_h}(t)\sin 2\omega_{cc}t$$

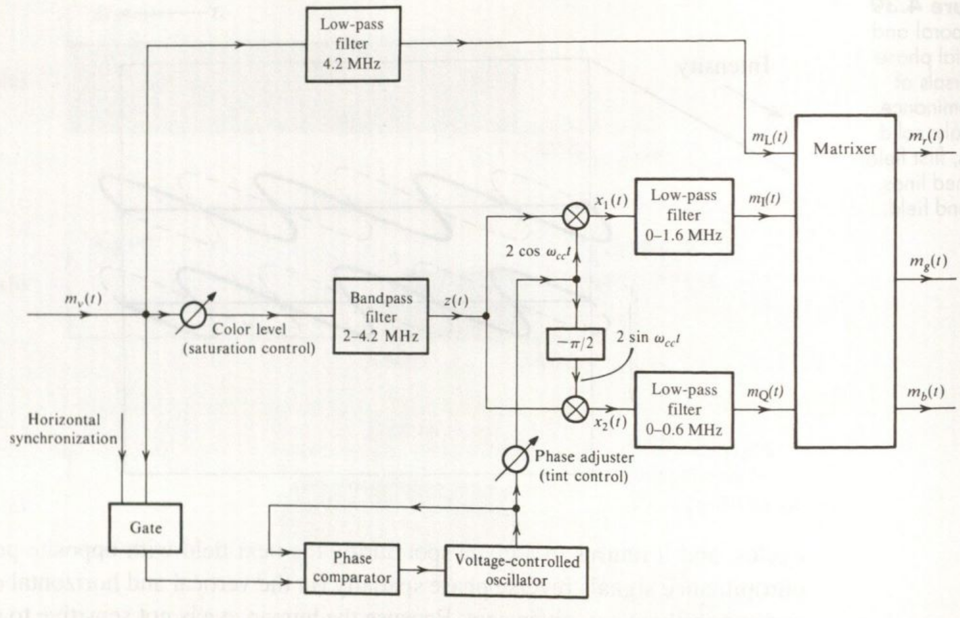
The double-frequency terms will be suppressed by the bandpass filter. In addition, the signal $2m_{LH}(t)\cos\omega_{cc}t$ will be invisible because of the frequency-interlacing effect. This is because the spectrum of this signal is the spectrum of $m_L(t)$ shifted to $\omega_{cc} = 227.5\omega_h$, and it will become invisible because of the frequency interlacing discussed earlier. Hence, the filter output of the 0 to 1.6 MHz filter yields $m_I(t)$. Similarly, the output of the 0 to 0.6 MHz filter* yields $m_Q(t)$. The three signals $m_L(t)$, $m_I(t)$, and $m_Q(t)$ are then matrixed to obtain $m_r(t)$, $m_g(t)$, and $m_b(t)$.

A color carrier is generated by using PLL. For this purpose we separate the color burst (Fig. 4.40) and apply it to a PLL whose output is the locally generated carrier that tracks the color burst. The phase of the locally generated carrier is adjustable. This is called **tint control**.

---

* This filter will suppress $m_{IH}(t)$, whose components lie in the range of 0.6 to 1.6 MHz.

Figure 4.40
Color television receiver.


Low-pass filter 4.2 MHz
$m_L(t)$
Matrixer
$m_r(t)$
$x_1(t)$
Low-pass filter 0–1.6 MHz
$m_I(t)$
2 cos $\omega_{cc}t$
$m_g(t)$
$m_v(t)$
Color level (saturation control)
Bandpass filter 2–4.2 MHz
$z(t)$
$-\pi/2$
2 sin $\omega_{cc}t$
Low-pass filter 0–0.6 MHz
$m_Q(t)$
$m_b(t)$
$x_2(t)$
Horizontal synchronization
Phase adjuster (tint control)
Gate
Phase comparator
Voltage-controlled oscillator

4.10 MATLAB EXERCISES

In this section, we provide MATLAB exercises to reinforce some of the basic concepts on amplitude modulations covered in earlier sections. Our examples illustrate the modulation and demodulation of DSB-SC, AM, SSB-SC, and QAM.

DSB-SC Modulation and Demodulation

The first MATLAB program, `triplesinc.m`, is to generate a signal that is (almost) strictly band-limited and consists of three different delayed version of the sinc signal:

$$m_2(t) = 2\,\mathrm{sinc}\left(\frac{2t}{T_a}\right) + \mathrm{sinc}\left(\frac{2t}{T_a}+1\right) + \mathrm{sinc}\left(\frac{2t}{T_a}-1\right)$$

```
% (triplesinc.m)
%   Baseband signal for AM
%   Usage m=triplesinc(t,Ta)
    function m=triplesinc(t,Ta)
%   t is the length of the signal
%   Ta is the parameter, equaling twice the delay
%

    sig_1=sinc(2*t/Ta);
    sig_2=sinc(2*t/Ta-1);
    sig_3=sinc(2*t/Ta+1);
    m=2*sig_1+sig_2+sig_3;
end
```

The DSB-SC signal can be generated with the MATLAB file `ExampleDSB.m`, which generates a DSB-SC signal for $t \in (-0.04, 0.04)$. The carrier frequency is 300 Hz. The original